

# *Crossbred cow adoption and its correlates: countable adoption specification search in Sri Lanka's small holder dairy sector*

Article

Accepted Version

Edirisinghe, J. C. and Holloway, G. J. ORCID:  
<https://orcid.org/0000-0002-2058-4504> (2015) Crossbred cow adoption and its correlates: countable adoption specification search in Sri Lanka's small holder dairy sector. *Agricultural Economics*, 46 (S1). pp. 13-28. ISSN 0169-5150 doi:  
<https://doi.org/10.1111/agec.12195> Available at  
<https://centaur.reading.ac.uk/39336/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1111/agec.12195>

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

## **CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

## **Countable Adoption Specification Search In Sri Lanka's Small-Holder Dairy Sector**

**Jagath C. Edirisinghe<sup>a</sup> and Garth J. Holloway<sup>b</sup>**

<sup>a</sup>Wayamba University of Sri Lanka, Makandura, Gonwila, Sri Lanka.

<sup>b</sup>Department of Agricultural and Food Economics, The University of Reading,  
PO Box 237, Earley Gate, Reading, RG6 6AR, United Kingdom.

Considerable specification choice confronts countable adoption investigations and there is need to measure, formally, the evidence in favor of competing formulations. This paper presents alternative countable adoption specifications – hitherto neglected in the agricultural-economics literature – and assesses formally their usefulness to practitioners. Reference to the left side of de Finetti's (1938) famous representation theorem motivates Bayesian unification of agricultural adoption studies and facilitates comparisons with conventional binary-choice specifications. Such comparisons have not previously been considered. The various formulations and the specific techniques are highlighted in an application to crossbred cow adoption in Sri Lanka's small-holder dairy sector (80 words).

**Keywords:** Count-data adoption, specification search, Bayesian unification of agricultural adoption studies, Sri Lankan small-holder dairy sector.

**Journal of Economic Literature classifications:** Q12 Micro Analysis of Farm Firms, Farm Households and Farm Input Markets; O33 Technological Change: Choices and Consequences; Diffusion; C11 Bayesian Analysis

---

Correspondence to: Jagath C Edirisinghe, Department of Agribusiness Management, Faculty of Agriculture and Plantation Management, Wayamba University of Sri Lanka, Makandura, Gonawila (NWP), Sri Lanka, Tel: 00 94 31 2299246 (office) 00 94 11 2963507 (Home) 00 94 71 3513883 (Mobile), E-mail: jagathed@yahoo.com.

## **1. Introduction**

In subsistence agricultural settings, crossbred cow adoption is a potent precipitator encouraging marketable surplus, stimulating entry into milk markets, and, often, alleviating the incessant burden of poverty. Crossbred cows arrive as counts. When new, improved technology avails itself in finite countable amounts, complications confront the investigator wishing to model the adoption decision. With the objective of understanding, better, the factors contributing to crossbred cow adoption, we compare and contrast the evidence in favor of prototypical count formulations. Despite some notable exceptions (see Holloway et al, 2000; Ahmed et al., 2002; and Abdulai and Huffman, 2005), specification choice relevant to countable adoption has been neglected in the agricultural-economics literature. Considerable scope therefore exists for examining formally the countable adoption process surrounding cross-bred cow adoption. Viewing the adoption decision under alternative statistical specification allows us to substantiate better inferences about adoption, the selection of various covariates among the total that avail themselves, and the validity of predictions made from the various model structures. We focus on a comprehensive set of characterizations in what we refer to, generically, as ‘count data models’ and we incorporate conditioning factors into the adoption process by extending the basic formulations into generalized-linear-model specifications.

The paper is organized as follows. In section two, we review key facets of the adoption literature, paying specific attention to the adoption of crossbreed cows. In section three we review, briefly, the sample setting and the data generation experiment. In section four we develop relevant empirical models. In section five we formalize one very important unifying feature of adoption studies derived explicitly from the present methodology, and argue its importance for future inquiry. Conclusions are offered in section six.

## **2. Motivation, objectives and modelling strategy**

Literature on adoption and diffusion of innovations is voluminous. Evidence abounds within the reviews by Feder et al. (1985), Besley and Case (1993), Sunding and Zilberman (2001) and, more

recently, Doss (2006). The coverage is heterogeneous. However, thematic developments identify commonalities including decision-making simultaneities surrounding the adoption process; predicting productivities between adopters and non-adopters; understanding adoption learning processes; conceptualizing social networking's adoption impacts; understanding neighborhood effects; and accounting for abstinence. Doss (2006) identifies neglected foci, including policy, institutions, markets and infrastructure and how the interaction of these factors can effect adoption of potentially beneficial technology and alleviate or mitigate poverty (Doss 2006).

### ***The 'binary choice adoption legacy'***

The overwhelming majority of studies reviewed by Feder et al. (1985), Besley and Case (1993), Sunding and Zilberman (2001) and Doss (2006) focus attentions on binary choice. Within binary-choice frameworks a set of standard procedures and econometric specifications arise and constitute an essential 'toolbox' available to agricultural economists. However, binary choice is not exclusively the *modus operandi* in agricultural technology adoption studies. Indeed many adoption processes, by their nature, lend themselves to more complex econometric specification. Livestock, in particular, avail in countable finite units and in the context of developing livestock-adoption methodology it is important to have available methodology for handling counts.

In the absence of counts, and the presence of cross-sectional data, investigations center on the binary choices confronting decision-makers (Besley and Case 1993). Within the binary-choice setting, the probit model is the most celebrated adoption model. Examples abound (see, for example, Misra, Carley et al. 1993; Foltz and Chang, 2002; Lapar and Ehui 2004; Koundouri, Nauges et al. 2006); closely followed in rank is the ordered probit model (see, for example, Lal 1999; Negatu and Parikh 1999; Giunta and Trivieri 2007), the multinomial probit (see, for example, Koop, 2003) and the binary and multinomial logit formulations (see, for example, Feder and Slade 1984; Staal, Baltenweck et al. 2002).

### ***Livestock innovations and subsistence household decision-making***

Livestock are of fundamental importance to a vast swath of humanity faced with subsistence

household decision-making (Kiwuwa et al, 1983). Indeed, one of the Consultative Groups of International Agricultural Research—the International Livestock Research Institute—focuses its attentions almost exclusively on the interface between human decision-makers and animal enterprise. If “the most important resource available to a subsistence household is the total amount of time that its members have available to spend in productive enterprises (Holloway and Ehui, 2001)” then the second-most important household resource is usually an animal input (Ehui et al., 1998).

Despite their fundamental importance, we have neither the literary legacy nor the methodological dexterity for dealing with adoption decisions surrounding animal counts or other countable, adoptable quantities. Despite the fundamental importance of animal agricultural activities, the vast majority of adoption studies focus on high yielding technologies within the crop sector (Abdulai and Huffman 2005). Research on adoption of technologies relating to animal inputs is limited and research relating to dairy-cow adoption is scant.

Like the crops sector, in which adoption decisions focus upon the introduction of high-yielding varieties, within the dairy sector adoption of crossbreed cattle are at issue. Evidence from one International-Livestock-Research-Institute experiment (see Holloway and Ehui (2002) for an overview) suggests that the adoption of high-yielding, intensively-managed European breeds for purposeful fertilization with grade indigenous stock spawns externality. Generally speaking, evidence (Holloway, barett and Ehui, 200a; Holloway and Ehui, 2001; and Holloway, Teklu and Ehui, 2008) suggests that crossbreeding is a potent policy option enabling formerly subsistence households to participate in emerging milk markets; has important implications for the food security of small farmers (Ahmed, Ehui et al. 2002); and, perhaps more importantly engenders statistically significant income gains from cow ownership (Nicholson et al., 2004).

### ***Experiment***

Given these antecedents, an experiment akin to the one enacted by The International Livestock Research Institute in Addis Ababa, is assessed in an altogether different climatic regime within Sri

Lanka (Edirisinghe, 2010). This experiment provides the basis for this study, the desire for improved understanding of the methodologies available to us or handling countable quantities and desire for unification of choices surrounding the countable adoption decision. In most of the cases, the model choice is dictated by the data availability and the research questions are addressed using frequentist procedures. The plethora of model choices often leads to an ‘embarrassment of riches’ leading to sometimes conflicting predictions about the adoption decision; the specific covariates deemed to affect the adoption decision; and, conditional on these covariates being chosen, their relative and absolute potency in policy prescription. The investigator, however, must have uncertainty given the breadth of options available for modelling the adoption decision. And the presence of uncertainty leads us to the inevitable question of models comparisons.

### ***Models comparisons***

Formal incorporation of countable decision-making into the conventional household-production framework (Singh, Squire and Strauss, 1986) arises in a latent ‘step-function model’ (Holloway, Barrett et al. 2000) developed specifically for the purpose of correlating the countable adoption quantities with other real-valued data generated from household decision-making. That contribution, like the contributions of Ahmed et al, (2002) and Abdulai and Huffman (2005) ignores potentially important information arising in the process of developing conjugate models for countable adoption. Conjugate models have the property that the prior and posterior descriptions of the parameter environment have the same functional form. Conjugate models are, necessarily, simplistic and are criticized because they over-simplify the complexities surrounding actual, countable adoption processes. Conjugate models, however, possess a very attractive property which makes them manifest within the modelling environment, especially when the investigator has uncertainty concerning model specification. To the best of our knowledge, within agricultural economics, these benefits appear to have been largely ignored to date.

### ***Conjugacy benefits***

When specification choice arises the investigator needs to compute the marginal density for the

data, which we signify as  $f(\mathbf{y})$ . A conjugate model makes  $f(\mathbf{y})$  available in closed form via direct calculation. When the model space is large, as it is in one of the settings considered within this paper, the ready availability of  $f(\mathbf{y})$  enables the investigator to compare and contrast, on formal probabilistic grounds, the most suitable model specification for the data. One purpose of this contribution is to illustrate this procedure in a natural-conjugate count-data setting. The natural-conjugate count-data models include, most notably, *the Poisson sampling model*, *the Geometric sampling model*, *the Negative-Binomial sampling model* and the relatively little used, though no less important, *Binomial sampling model*. In the cases of two of these models (the Negative-Binomial and the Binomial sampling models), the mathematical formulation includes an assigned integer value, making comparisons across a large number of potential models straightforward by virtue of the availability of  $f(\mathbf{y})$ . A second purpose of this contribution is to illustrate this search procedure.

### ***Integer search***

We illustrate the vast benefits available to the researcher—measured in formal probability gains—of correctly assigning integer values to the Negative-Binomial and Binomial sampling models. Notwithstanding, the aforementioned benefits, the conjugate settings (Poisson, Geometric, Negative-Binomial and Binomial sampling models) are restrictive. And they are particularly restrictive in the search for prescriptions concerning the policies we should enact in order to promote adoption. They are restrictive within the policy domain because they are restrictive primarily in understanding the fundamental intricacies surrounding the adoption decisions undertaken by subsistence small-holders. These issues, it is worth emphasizing, are the heart of the literature reviewed in the flagship contributions (Feder et al, 1985; Besley and Case, 1993; Sunding and Zilberman, 2001; and Doss, 2006). For the purpose of understanding the nuances surrounding the adoption decision we need a wealth of additional information, or, at the very least, the usual set of approximations to which investigations default, which is covariates. A third purpose of this contribution is to illustrate the incorporation of covariate information into the count-data adoption framework.



### *Covariate search*

When covariate information is available, we incorporate the covariates by way of a generalized-linear model. Such models have become commonplace in related settings in which a conjugate model need default to a generalized linear framework in order to include covariate information. The potential benefits of including covariate-information, however, are necessarily complicated by costs of having no longer available in closed form the desired quantity  $f(\mathbf{y})$  for model choice. In general, only for the Normal-Linear-Regression and Log-Normal-Linear Regression setups does the model retain its conjugate properties when parameters are replaced by covariates and corresponding regression coefficients. The non-availability of  $f(\mathbf{y})$  directly in closed form complicates analysis requiring us to produce estimates of  $f(\mathbf{y})$  and necessarily restricts the sample space. This situation is not novel. For example, previously within this journal, Hattam, Lacombe and Holloway (2012) model time-to-adopt organic practices and consider model choice across conjugate settings which relinquish conjugacy in the presence of covariates.

The significance for model selection of the absence of conjugacy is that it confines attentions in the insurmountable model space with  $K$  covariates under consideration in which evaluations of all  $2^K$  possible models are sought. Presently, we have four model specifications under consideration (Poisson, Geometric, Negative Binomial and Binomial); an integer search across two of these specifications (the Negative Binomial and Binomial models); and model comparisons within each specification amounting to  $2^K$  where  $K$  is potentially large. The Cartesian product (specifications  $\otimes$  integer values  $\otimes$  covariates) defining the model space is too large for us to make definitive statements across the entire space. However, we argue, through a suitable use of the conjugate information *prior to the addition of covariates*, along with a restricted covariate search conditioned by the results of model search under conjugacy, an approximate ‘preferred’ specification emerges. Thus, our fourth, and final, objective is to indicate how judicious choice of conjugate and non-conjugate results can be combined to motivate, a ‘preferred’ specification suited to the adoption process. We enact these demonstrations following presentation of details concerning the Sri Lankan

sample setting and a description of the data.

### **3. The Sri Lankan sample setting and the data**

There are approximately 1.5 million cattle spread throughout Sri Lanka, with areal concentrations due to cultural, market and agro-climatic factors favoring specific breeds, specific production practices, and specific markets and marketing activities. Smallholders account for approximately 96 percent of the total cattle-rearing population according to the most recent agricultural census (Department of Census and Statistics, 2002). The overall penetration of exotic breeds is considered to be ‘low’ (Department of Census and Statistics, 2002).

#### ***Sampling strategy***

We enact sampling taking into consideration variation in agro-climatic factors that are thought to impact adoption propensity. Sri Lanka has three main agro-ecological regions, which, for want of better nomenclature, we refer to, respectively, as the ‘dry,’ ‘wet,’ and ‘intermediate’ zones. Sri Lankan agriculture is also classified with respect to ‘elevation.’ We employ the nomenclature ‘upper,’ ‘middle,’ and ‘lower’ elevations in order to distinguish between elevation levels. Both the wet and intermediate zones are observed in each of the three elevation levels but the dry zone is mostly limited to the lower elevation. In addition to agro-ecological zones, sampling is also motivated by the delicate balance between ‘costs’ and ‘benefits.’ Needing to restrict costs of data acquisition we target representative dairying districts that feature each of the agro-climatic regions thought to influence production (Department of Census and Statistics, 2002). And within these eight districts we sample as expansively as resources permit, with coverage amounting to 401 households. The total sample size is distributed among the districts based on the desire for an approximate proportionate representation within each district, as reported in the most recent agricultural census (Department of Census and Statistics, 2002). Within Sri Lanka, political administrative units are ordered in hierarchical fashion, with the smallest administrative unit being known as a ‘Grama Niladhari’ division. We use each of the Grama Niladhari veterinary officer’s lists of farmers in order to randomly select households from within each of the eight divisions

comprising the sample.

### ***Responses***

Figure 1 depicts the distribution of the response variable—the total number of crossbred cows within each household—across the 401 households. We make reference to the observation that the mean number of crossbred cows is 1.77 and that the variance is small, being, approximately 2.45. We emphasize that the sample estimates of mean and variance do not coincide, which is a condition assumed by the benchmark Poisson model. In fact the variance exceeds the mean by approximately 39 percent and, alone, posits grounds to consider alternative model formulations. (Insert figure 1 about here.)

### ***Agro-ecological zone distribution***

The sample size is  $N = 401$ . The minimum count is zero, which represents a small proportion of the observations; and the maximum count is 11, which is observed in one single household. One issue of considerable interest across the sample is the relative difference, if any, in the distribution of the counts across the three climatic zones. The distribution of the sample across the three zones is as follows. There are 152 dry-zone respondents; there are 93 wet-zone respondents; and there are 156 intermediate-zone respondents. Figure 2 depicts the climatic-zone conditional distribution of the cow counts across the sample along with the sample total distribution across counts. Some minor differences emerge, although the overall conditional and marginal count distributions are quite similar. (Insert figure 2 about here.)

### ***Covariate justification and summary statistics***

Data are collected from face-to-face interviews with the help of a structured questionnaire. The questionnaire includes data on the number of crossbred animals employed by the household; demographic characteristics; the human-, social-, financial- and physical-capital status of the household; perceptions about crossbred cow technology; elevation levels; and transactions costs. The selected covariates are chosen based on three factors. One factor is the background information contained in the relevant literature (surveyed by Feder et al., 1985; Besley and Case, 1993; Sunding

and Zilberman, 2001; and Doss, 2006); a second factor is the lead author's previous site experience and familiarity with Sri Lankan farming systems; and the third factor is casual, contemporaneous empirical evidence gleaned intermittently during primary data collection.

Our focus is on the few selected covariates. We include the agro-ecological zones dry, wet and intermediate because we expect to observe different practices and perception of cross-bred cow technology within each climatic zone and, therefore, because zone difference is expected to explain some variation in adoption practices, we wish to control for zone difference. We include 'time to market' to reflect market access and developments in infrastructure and because time-to-market features as a prominent covariate in each of the milk-market precipitation studies produced by Holloway and co-authors (see, for example, Holloway and Ehui, 2001). We include income levels of the farmers in question because wealth, in general, but 'Income,' in particular, is likely to affect aversion to risk. The variable 'Land' is included because it is thought that congestion effects may bottleneck the adoption process and because there is uncertainty generally within the literature surrounding the size-versus-efficiency conundrum (see, for example, Barrett (1996) and the references cited there). The variable 'Cattle Shed Dummy' is included in order to identify those individuals having access to superior housing conditions for cross-bred animals; this variable is likely highly correlated with wealth and status effects in the adoption decision. The tri-chotomous variable 'Perception I' is included in order to represent overall indications of the farmer's view of intensive crossbred farming practices. The variable 'Perception II' is included in order to condition perceptions about costs factors. The effect of institutional developments is taken into account by including the membership status with a milk society, as depicted by the variable 'Society. Feder (1985) states that "Farmers' technology choices are based on their subjective probabilities and hence, on their exposure to information regarding the new technology". We include 'Education' as a (somewhat crude) projector of the ability of the farmer to process information. We hypothesize that access to information is reflected by the level of contact with the agents and use the variable 'Extension' following a considerable precedent (see, in particular, Frisvold et al, 2001; Holloway

and Ehui, 2001; Dinar and Keynan, 2001; and Hanson and Just, 2001).

### ***Summary***

In summary, we believe that cross-bred cow adoption is affected by (a) the agro-ecological climatic zone in which decisions are made, (b) proximity to market, (c) wealth effects inclusive of income, access to land, and cattle sheds, (d) perceptions about crossbreeding and crossbreeding production practices, (e) social status, social networks and neighborhood effects, and by (f) the intellectual capital accumulators, experience education and experience. These effects are accounted for through reference to the (numbered) covariates depicted in table 1. (Insert table 1 about here.)

With reference to the covariates, it is insightful to consider inter-component differences across the adopting and non-adopting sample constituents as well as differences across agro-ecological zones. Differences among the agro-ecological zones are depicted within table 2. Within the dry, wet and intermediate zones there are 152, 93 and 156 respondents, respectively; and within the non-adopting and adopting classifications there are 51 and 350 respondents, respectively. The disposition of the covariates is mostly homogeneous, with a few exceptions. Some small but potentially significant differences emerge. These differences raise scope for nuanced empirical investigation about the factors precipitating countable cow adoption among the Sri-Lankan smallholders. (Insert tables 2 and 3 about here.)

## **4. Empirical models and results**

Given the data-generating experiment, we outline, within this section, our inferential methodology. Relegating details to the appendix to the manuscript, contained at the end of the text; we also provide additional inferential detail in an expanded online version of the appendix. We focus discussion on the presentation of key results affecting model choice, covariate selection and various other motivations that guide and direct enquiry. The spirit of enquiry lies in orchestrating analysis exhaustively prior to considering covariates. As usual, in this setting, all inferential questions are guided by formal probability considerations.

### ***Notation***

Preliminary investigation of the four, conjugate settings evolves as follows. Using  $\mathbf{y} \equiv (y_1, y_2, \dots, y_N)'$  to denote the observed counts across the ( $N = 401$ ) households, we use  $f(\cdot)$  to denote, generically, an appropriate probability density function; use  $f(\mathbf{y}|\boldsymbol{\theta})$  to denote the conditional distribution for the data given the parameter, which may be vector-valued, say  $\boldsymbol{\theta} \equiv (\theta_1, \theta_2, \dots, \theta_K)'$ ; and use  $f(\boldsymbol{\theta})$  to denote the distribution for the parameters in the absence of data, which we refer to, henceforth, as the ‘prior probability density function (pdf).’ In what follows, and in most usual circumstances, we are primarily interested in the scales and locations of the marginal pdf’s for the parameters after examining the data and which we refer to henceforth as the ‘posterior probability density function (pdf).’ The joint distribution is  $f(\boldsymbol{\theta}|\mathbf{y})$  and the marginal distributions of each (scalar) parameter will be referred to accordingly as  $f(\theta_1|\mathbf{y})$ ,  $f(\theta_2|\mathbf{y})$ , ...,  $f(\theta_K|\mathbf{y})$ . For most enquiries, the derivation of the densities  $f(\theta_1|\mathbf{y})$ ,  $f(\theta_2|\mathbf{y})$ , ...,  $f(\theta_K|\mathbf{y})$  identifies the end of the investigation. Presently, however, interests center on model comparisons. For this reason, the marginal density of the data’ also termed ‘the evidence’ also termed ‘the marginal likelihood’ makes the quantity  $f(\mathbf{y})$ , of over-arching importance. In most investigations, densities,  $f(\theta_1|\mathbf{y})$ ,  $f(\theta_2|\mathbf{y})$ , ...,  $f(\theta_K|\mathbf{y})$  are also conditional upon investigator choice. Where there is uncertainty associated with such choice, ‘model realizations’ need to be evaluated as part of the inferential process. For example, we may be interested in the densities  $f(\theta_1|\mathbf{y})$ ,  $f(\theta_2|\mathbf{y})$ , ...,  $f(\theta_K|\mathbf{y})$  in the presence of covariates, which we signify  $\mathbf{X}$ , defining a matrix of observations (of dimension  $N \times K$ ), containing rows  $\mathbf{x}_1$ ,  $\mathbf{x}_2$ , ...,  $\mathbf{x}_N$ , wherein  $\mathbf{x}_1 \equiv (x_{11}, x_{12}, \dots, x_{1K})$ ,  $\mathbf{x}_2 \equiv (x_{21}, x_{22}, \dots, x_{2K})$ , ...,  $\mathbf{x}_N \equiv (x_{N1}, x_{N2}, \dots, x_{NK})$ , and  $\mathbf{X} \equiv (\mathbf{x}_1', \mathbf{x}_2', \dots, \mathbf{x}_N')'$ . In this context, we may write,  $f(\mathbf{y}|\boldsymbol{\theta}, \mathbf{X})$  to denote the conditional distribution for the data given the parameters and the observed covariate information; use  $f(\boldsymbol{\theta}|\mathbf{y}, \mathbf{X})$  to denote the posterior pdf for the parameters given the data; and use  $f(\mathbf{y}|\mathbf{X})$  to denote the marginal distribution for the response data conditioned by the covariates.

### ***Specification search***

Our task is to evaluate  $f(\mathbf{y})$  and  $f(\mathbf{y}|\mathbf{X})$ . This task is further complicated by the fact that the matrix  $\mathbf{X}$

involves, itself, a choice. When there is a single column within  $\mathbf{X}$  ( $K = 1$ ), we have available one choice with either of two mutually-exclusive outcomes. We may include  $\mathbf{X}$  or exclude  $\mathbf{X}$  from the proceedings. When there are two columns in  $\mathbf{X}$  ( $K=2$ ), there are four outcomes. And when the columns in  $\mathbf{X}$  number three ( $K=3$ ), there are eight possible realizations. Considering that  $2^1$ ,  $2^2$  and  $2^3$  defines the total number of outcomes within each setting, we wish to consider the total number of models available,  $M$ , which is  $2^K$ . Using  $\mathcal{M}^k$  to denote the choice under action  $k$ , where  $k = 1, 2, \dots, M$ ; we can write, in turn,  $f(\boldsymbol{\theta}|\mathcal{M}^k)$ ,  $f(\mathbf{y}|\boldsymbol{\theta}, \mathcal{M}^k)$ ,  $f(\boldsymbol{\theta}|\mathbf{y}, \mathcal{M}^k)$  and  $f(\mathbf{y}|\mathcal{M}^k)$  to denote, respectively, *the prior pdf, likelihood for the data, the posterior pdf and the evidence*, all of which are conditional on the realizations of the choice made by the investigator. One important task of this paper is to evaluate the evidence in favor of competing models, indexed  $\mathcal{M}^k$ ,  $k = 1, 2, \dots, M$ . Conceptually, we seek to evaluate the ‘mixtures’

$$f(\mathbf{y}|\mathcal{M}^k) = \int f(\mathbf{y}|\boldsymbol{\theta}, \mathcal{M}^k) f(\boldsymbol{\theta}|\mathcal{M}^k) d\boldsymbol{\theta}, \quad k = 1, 2, \dots, M. \quad (1)$$

Remaining attentions are devoted to evaluations of (1), within which, henceforth, in order to avoid notational clutter, we suppress dependency on the model label,  $\mathcal{M}^k$ .

### ***Search***

Focusing attentions on  $f(\boldsymbol{\theta})$ ,  $f(\mathbf{y}|\boldsymbol{\theta})$ ,  $f(\boldsymbol{\theta}|\mathbf{y})$  and  $f(\mathbf{y})$  under the respective models specifications, we emphasize that the *Poisson*, *Geometric*, *Negative Binomial* and *Binomial* model formulations right sides of equation (1) are available in closed form. In particular, the marginal likelihoods, for a particular model in the Poisson, Geometric, Negative Binomial and Binomial settings are known to be, respectively, Poisson-Gamma, Negative-Binomial-Beta subject to a parameter restriction; Negative-Binomial-Beta; and Binomial-Beta. This feature of the analysis is particularly noteworthy in the last two settings in which further choice avails itself. Unlike the Poisson and the Geometric models, the Negative-Binomial and the Binomial models are count-data models containing a second, integer-valued parameter. While the Negative Binomial and Binomial models remain conjugate, it is straight-forward to evaluate both of them across a range of values for these integer-

valued parameters. Because it is then possible, due to their inherent conjugacy, to compute the left-hand side of (1), by direct computation, it is possible to easily evaluate the evidence in favor of any single integer-valued specification. Consequently, inferences about the adoption process obtained from a specific model can be assigned their formal probability weight in a model averaged setting. Increasingly, model averaging is appearing in a plethora of applied investigations as an appropriate procedure for dealing with uncertainty within the high-dimensional model space. A good discussion of this procedure now appears in graduate textbook treatments, and the reader is directed to Koop (2003) and Koop, Poirier and Tobias (2007) for details. At this point two questions beckon. First, to what extent do the foundational conjugate count models explain the data generating environment? Second, will there be improvement, if at all, once covariate information is included? We answer these questions, in the following sequential manner.

### ***Preliminary conjugate analysis***

Given the data (figure 1) and some sensitivities about how the propensities may change across climatic zones (figure 2), we first consider straight-forward conjugate analysis of, respectively, Poisson, Geometric, Negative Binomial and Binomial sampling models. In the latter two cases, one must select an integer value for the negative-Binomial parameter ' $r_o$ ' which is defined with respect to the Negative Binomial pdf, presented in the appendix. In the Binomial setting, we must also adopt a value for the parameter ' $n_{oi}$ ,' which is defined with respect to the Binomial pdf, presented in the appendix. The maximum number of cows owned within the sample is 11. It would not seem unreasonable, therefore, to consider this household, somewhat an outlier (figure 1), as being 'successful' in each of 11 consecutive trials. Such an assumption would mandate  $n_{oi} = 11$  as the lower limit, but other assumptions cohere. In the Negative Binomial model the assumption  $r_o = 1$  leads to the Geometric distribution. For reasons which we soon justify, formally, we select  $r_o = 7$  and we assign  $n_{oi} = n_o = 139$  to each household. Given these specializations, we are now in a position to consider inferences under each respective specification. The results of case-specific analysis (please refer to the supplemental appendix) are summarized in figures 3 and 4. (Insert



figures 3 and 4 about here.) In figure 3 we report, for each count, the implied probabilities distributed across the sample, including the empirical conditional probability derived from the data and the *Poisson*, *Geometric*, *Negative Binomial* and *Binomial* probabilities. By and large, the four models do a good job of predicting the counts. The often encountered ‘over-inflation’ of zeros that arises in the Poisson model is evident, but is insubstantial. Notwithstanding this feature, all models do a relatively poor job of picking up the second element of counts, the value 1, which is the modal value within the sample. Taking these reports, we are able then to transform the results into a figure reporting the relationship between the observed and predicted probabilities within the data. We note, that correlations between the observed and predicted quantities are quite high, and are dominated—not exclusively, but collectively—by the Poisson, Negative-Binomial and Binomial models; the Geometric model is adjudged, by this criterion, to perform rather poorly.

### ***Exchangeability***

We now consider the problem of choosing and selecting across the various models, and a small digression, inclusive of some significant output contained within the supplemental appendix, is in order. We motivate that the different agro-ecological zones give rise to distinctly different expectations about adoption propensity, and we note that the knowledge of existence within each zone is datum. The question is just how significant are these differences? We are able to answer this question by appealing to a notion which is frequently employed to justify Bayesian inference, but which is largely untapped – to our knowledge – in applications. This notion is that of ‘*exchangeability*.’ Exchangeability was promulgated by Johnson *circa* 1920 (Johnson, 1924) and subsequently exploited in a ground-breaking contribution by de Finetti (1938). The consequences of de Finetti’s theorem—the so-called ‘representation theorem’—are profound for model building, inference and selection (Bernardo, 1996). Kreps (1988) opines that de Finetti’s theorem is the ‘fundamental theorem of statistical inference.’ And Poirier (2011) concludes that de Finetti’s theorem places the entwined notions of exchangeability, subjectivity and model building in their appropriate inferential contexts. Its relevance within applications—hitherto neglected in agricultural

economics—is fundamentally important, as follows. By sequentially updating throughout the sample space we are able to gauge whether the exchangeability condition is violated and whether modification is called for. Specifically, wishing to evaluate the assumed belief that the propensities to sample from the three distinct agro-ecological zones are exchangeable, we evaluate the sequential updating of the posterior expectations about the adoption propensities (the likelihood of adopting a given number of crossbreed cows) and consider the outputs under alternative sequencing of the data. *If the data are in fact exchangeable, then the posterior inference is unaffected by the order in which the data are processed.* By reordering the observations specific to various regimes we can observe a difference—if any—that is available from re-ordering. *If the observations within the sample are indeed exchangeable then we should have no concern about the agro-climatic locations of the individual households and a single-sample processing of the data is applicable. If we find to the contrary, then modification is called for.*

We consider the relevant re-sequencing and sequential updating within the sample within the graphics which appear in the supplement (see , in particular, Supplemental Figures S.i.j,  $i = 1, 2, 3, 4$ ,  $j = 1, 2$ ). The figures indicate that, clearly, within the separate evaluations of the Poisson, Geometric, Negative Binomial and Binomial models, exchangeability of agro-ecological zones is violated. Modification is necessary. Either the inclusion of binary, climatic-specific indicators (fixed effects) or extension to hierarchical formulation (the Bayesian analogue of random effects) is called for. (We elaborate on considerations of exchangeability in the supplemental material available to the reader online.)

### ***Integer search strategy***

We now turn to the question about the evidence in favor of each competing formulation and note that considerable choice is available. Prior to considering covariates, we compute the model marginal likelihoods corresponding to the Poisson, Geometric, Negative Binomial and Binomial models, where the Negative-Binomial specification is evaluated under an iterated search across the parameter ' $r_o$ ,' and the Binomial specification is evaluated under an iterated search for the

parameter ‘ $n_{oi}$ .’ Figure 5 reports the results of this search. (Insert figure 5 about here). The results are quite revealing. First, no single model dominates over all ranges within the search. Second, each of the three alternate models Poisson, negative Binomial and Binomial dominate the Geometric model. Third, the Negative Binomial model has an optimal (‘optimal’ in the sense that it maximizes the corresponding log marginal likelihood) value of the parameter ‘ $r_o$ ’ which is  $r_o = 7$ , and thus the motivation for the aforementioned choice is justified. Fourth, the Binomial model has an optimal value for its ‘trials’ parameter, which is  $n_o = 139$  and its reason for pre-assignment is revealed.

Given the marginal likelihood values reported in the figure, it is relatively straightforward to compute the probabilities in favor of any particular model. These reports are presented in figure 6. (Insert figure 6 about here.) We note that the maximum probability is around .166 and that the bulk of the probability space is dominated by about ten models. We note also that each of these ten models relates to different parameterizations of the Negative Binomial model. That is, in the absence of covariate information the Negative-Binomial specification, variously parameterized, dominates the sample space.

### ***Covariate search strategy***

Taking as datum the integer values  $r_o = 7$  and  $n_{oi} \equiv n_o = 139$ , we next consider the covariates predicting the counts. Given the 13 covariates under consideration (please refer to table 1 and the numbering of the covariates), taking as datum the agro-ecological-specific dummy variables consider covariate selection across the remaining 10 covariates. A total of  $S = 2^K = 2^{10} = 1024$  models must be evaluated and these evaluations must be executed within each respective formulation. Thus, a total of  $S = 4096$  models must be evaluated. The totality of models defining the sample space is of course, considerably larger if we were to repeat the integer search across all 1024 covariate specifications. This computation is infeasible due to the fact that we need estimate, rather than calculate,  $f(\mathbf{y})$ , for each specific formulation.

We do not attempt to tackle this problem comprehensively. However, using the knowledge

acquired from the conjugate settings, with  $r_o = 7$  and  $n_o = 139$  performing well for the Negative-Binomial and Binomial models, respectively, we consider four searches within the vicinity of these restrictions in an effort to gain a better understanding about the likelihoods of ‘optimal’ outcomes throughout the covariate sample space. The four exercises are, first, a search across a reduced-design matrix consisting of the zone-specific dummy variables and the Negative-Binomial parameter values permitted to vary from  $r_o = 1$  to  $r_o=100$  (100 passes); the same experiment but with the full column design of the covariate matrix (100 passes); a search across the reduced-design matrix with the Binomial parameter values permitted to vary from  $n_o = 11$  to  $n_o = 210$  (200 passes); the same experiment but with the full column design of the covariate matrix (200 passes). In each case, we find that the log-marginal likelihood values are quite stable after a few thousand iterations, but we execute under the conservative procedure of sampling for  $S = 20,000$  iterations discarding the first 10,000 observations and processing the second 10,000 observations. Some interesting patterns emerge. These patterns are portrayed in figure 7. (Insert figure 7 about here.) The results for the restricted covariate matrix experiments are clearly inferior to the results reported for the unrestricted cases and appear to be bounded by the reports for the conjugate models.

### ***Preferred models specifications***

With this covariate-search information at hand, we combined findings with the results of the 1024 models run across the four count specifications with parameters ‘ $r_o$ ’ and ‘ $n_o$ ’ set at 7 and 139, respectively, and consider the ‘optimal’ configurations across each profile. For the purpose of permitting one final comparison, we permitted the simulations of each of the four models to run for a few days using randomly selected covariate specifications and permitting random selections of the integers between for the Negative Binomial and Binomial formulations. And we determine that respective full-covariate generalized-linear-model formulations marginal likelihood estimates are quite comparable to the random search values. The results of the estimations are reported in table 4. (Insert table 4 about here.) Within table 4 we observe a fairly heterogeneous profile of ‘significance’ across the four primary specifications. And we note that, at least according to the

marginal likelihood reports corresponding to the counts, that there is relatively close proximity between the Poisson, Negative Binomial and Binomial generalized-linear-model specifications. The Geometric generalized linear model specification—as it did within the conjugate setting (see, in particular, figure 5)—again performs poorly, at least as adjudged by the marginal-likelihood criterion. In addition to the count-data-generalized-linear model specifications we include, in the left-most region of the table, reports derived from a standard probit specification of the ‘adopt-versus-non-adopt’ binary outcome. A conventional investigation would most likely enact probit estimation when confronted with the counts, treating the counts  $y_i = 0$ , for  $i = 1, 2, \dots, N$ , signifying ‘non-adoption’ and treating the counts,  $y_i = 1, 2, \dots$ , for  $i = 1, 2, \dots, N$ , signifying ‘adoption.’ Until, presently, comparisons between counts and binary outcomes have been thought to be ill-considered due to the fact that the marginal density for the data,  $f(\mathbf{y})$ , pertains to a given set of observables. This is problematic within the present context due to the desire for understanding of the performance of conventional binary choice specifications given the binary-choice legacy in agricultural adoption studies. We demonstrate, in the next section, how one is able to circumvent this problem, specifically with reference to Bayes theorem and de Finetti’s (1938) identity.

## 5. Bayesian Unification of Agricultural Adoption Studies

It is noteworthy and emphasized that the majority of adoption studies—within agricultural economics and elsewhere—view adoption processes as the outcome of a binary-choice experiment. Within binary choice frameworks, the *modus operandi* are the familiar probit and logit specifications. Indeed, the probit specification is considered “the most celebrated binary choice specification (Koop, Poirier and Tobias, 2007).” Sometimes ordinal and categorical response specifications such as the ordered probit and the multinomial probit and logit specifications are appropriate. This contribution has added the count-data models specifications Poisson, Geometric, Negative Binomial and Binomial and generalized-linear extensions.

Given the rich heritage devoted by agricultural economists to binary choice it becomes useful, pedagogic and insightful to consider unification of the results of multifarious specifications—non-

binary choice—into a common whole. Here, we consider, explicitly, the problem of unifying agricultural-adoption studies and provide an example about how such unification proceeds. Our goal is to compute a single over-arching metric, which is capable of evaluating relevant sample information enabling a ‘preferred’ specification to emerge with reference to the adoption threshold. A solution is available, once again, with reference to the left-side of de Finetti’s famous 1938 identity, which is the quantity  $f(\mathbf{y})$ . As the reader, no doubt, observes, the recurrent dependence on,  $f(\mathbf{y})$ , gives this over-arching measure profound status, to which we assign, the alternative interpretation, ‘the statistical footprint of the data.’ For the purpose of evaluating  $f(\mathbf{y})$  presently, a hierarchical representation, similar in spirit to Albert and Chib (1993) and, similar in decomposition and rearrangement to Chib (1995), prove fruitful.

Consider the counts, once again, but suppose, specifically that the counts are realizations of some latent-variable process, say,  $N_{0i} = 0, 1, 2, \dots$ , relevant to household ‘i.’ Assume that this process gives rise to the observed, binary outcomes,  $y_i = 0$ , if  $N_{0i} = 0$ , and  $y_i = 1$  if  $N_{0i} \neq 0$ . Following Koop, Poirier and Tobias (2007), form a binary indicator density for the joint distribution of  $\mathbf{y}$  and  $\mathbf{N}_0 \equiv (N_{01}, N_{02}, \dots, N_{0N})'$  and following Chib (1995) derive the evidence in favor of the count-data model, from which the  $\mathbf{N}_0$  are observed, using the ‘basic marginal likelihood identity,’

$$f(\mathbf{y}) = f(\mathbf{y}|\mathbf{N}_0) \times f(\mathbf{N}_0|\boldsymbol{\theta}) \times f(\boldsymbol{\theta}) \div f(\boldsymbol{\theta}|\mathbf{N}_0, \mathbf{y}) \div f(\mathbf{N}_0|\mathbf{y}). \quad (2)$$

Here,  $f(\mathbf{y})$  denotes the marginal likelihood for the ‘binary-data-transformed’ counts. Within the numerator, the density  $f(\mathbf{y}|\mathbf{N}_0)$  denotes the indicator function depicting the binary outcomes in the transformation of the  $N_{0i}$ ’s to the  $y_i$ ’s. That is  $f(\mathbf{y}|\mathbf{N}_0)$  satisfies

$$f(\mathbf{y}|\mathbf{N}_0) = \prod_{i=1}^N I(y_i = 0, N_{0i} = 0) + I(y_i = 1, N_{0i} \neq 0). \quad (3)$$

The density  $f(\mathbf{N}_0|\boldsymbol{\theta})$  denotes the count-data sampling density. And the density  $f(\boldsymbol{\theta})$  denotes the prior pdf for the unknown  $\boldsymbol{\theta}$ . With respect to the denominator,  $f(\boldsymbol{\theta}|\mathbf{N}_0, \mathbf{y})$  denotes the posterior pdf for the unknown  $\boldsymbol{\theta}$  upon observing the counts (and, hence, the binary outcome). And the density  $f(\mathbf{N}_0|\mathbf{y})$  denotes the marginalized distribution for the counts, conditional on the binary outcomes. This density is simply the dichotomized pdf for the counts, reducing their probability space to the binary

outcomes. That is, for  $i = 1, 2, \dots, N$ ,  $N_{0i} = 0$ , if  $y_i = 0$ , and  $N_{0i} \neq 0$  if  $y_i = 1$ . Within conjugate settings we can compute directly  $\wp(N_{0i} = 0|y_i = 0)$  and  $\wp(N_{0i} \neq 0|y_i = 1) = \sum_{j=1}^{\max(N_{0i})} \wp(N_{0i} = j|y_i = 1)$ . In non-conjugate settings we estimate these probabilities via the usual *Monte Carlo method*, averaging across the draws within the sampling algorithm. Thus, despite  $f(N_{0i}|y_i)$ , having complex form, in principle, an estimate, say,

$$f(N_{0i}|y_i) \cong G^{-1} \sum_{g=1}^G f(N_{0i}|y_i, \theta^{(g)}) , \quad (4)$$

is available. The question of preference between the conventional probit specification and the non-conventional generalized-linear-count-model specifications can then be determined with reference to the estimates of  $f(y_i)$ , lodged at the base of table 4. A striking result emerges. This result is that the various count-data specification dominate the conventional probit specification within the binary-choice context. Moreover, within the preferred count-data specializations for predicting the binary outcomes, the Geometric model—performing disparagingly in conjugate and count-based generalized-linear-model settings—is now dominant. The result is striking and raises considerable scope for further nuances unifying agricultural adoption studies.

## 6. Concluding comments

This study considers countable adoption specification search in application to household dairy production. A rich set of Sri-Lankan small-holder dairy data are evaluated. The data are of interest in and of themselves but are potentially more interesting in light of our search for a plausible count-data model specification. In the context of the counts, some additional problems arise that are non-existent in more conventional adoption analyses. This paper has demonstrated how these problems can be addressed formally in a common unifying framework; has developed methodology for model selection and covariate comparison; and, importantly, assessed the potential gains from pooling the results from alternative specifications within a unifying whole. We have also proposed a direction for additional research with the aim of comparing, hierarchically, a greater plethora of specifications, some of which are novel and others which the traditional agricultural-economic contributions have produced. Work remains in further formalizing this unification. Until then the

results presented are subject to the usual caveat. This caveat is that an over-arching, hitherto neglected, but preferred model, may exist.

## **Acknowledgements**

We are grateful to three anonymous referees, *The Editor*, and to our friend and colleague Donald J. Lacombe, for numerous suggestions and constructive comments. Full responsibility for remaining error is ours.

## **Appendix**

This appendix subdivides into sections, including, background, notation, density development, conjugate models specification and non-conjugate models specialization. We facilitate implementation by amending well-known results in three, small, but important, ways. First, we extend the four foundations in counts, which are the Poisson, Geometric, Negative-Binomial and Binomial models to generalized-linear specializations. Second, we determine from an extensive set of models comparisons the preferred generalized-linear framework in each case. Third, we modify the basic marginal-likelihood identity outlined in Chib (1995) and extended in Chib and Jeliazkov (2001) in order to facilitate selection. In addition to Zellner (1971), Lindley and Smith (1972) and Chib (1995), some familiarity with conjugate developments in count-data models at a level of dexterity of, say, Mood et al (1974) or, alternatively, Poirier (1995), is advantageous. We assume familiarity with regression techniques in Markov Chain Monte Carlo systems as appears, for example, in Koop (2003) and Koop, Poirier and Tobias (2007).

### **A.1 Notation**

By way of notation we use lower-case Greek and Roman numerals to reference scalar quantities, use emboldened lower-case symbols to reference vectors and use emboldened upper-case symbols to reference matrix quantities. Thus, let  $\boldsymbol{\theta} \equiv (\theta_1, \theta_2, \dots, \theta_N)'$  denote a vector of parameters of interest, where  $'$  denotes the ‘transpose’ of the column vector  $\boldsymbol{\theta}$ ;  $\pi(\boldsymbol{\theta})$  denotes the prior probability density function (*pdf*) for  $\boldsymbol{\theta}$ ; and  $\pi(\boldsymbol{\theta}|\mathbf{y})$  the posterior *pdf* for  $\boldsymbol{\theta}$ ; where  $\mathbf{y} \equiv (y_1, y_2, \dots, y_N)'$  denotes data. Frequently, we reference the data generating model  $f(\mathbf{y}|\boldsymbol{\theta})$ , which is the likelihood function when



viewed as a function of  $\theta$  conditioned by  $y$ ; and, sometimes, make use of variants of the  $f(\cdot|\cdot)$  notation in order to reference particular probability density functions. Occasionally we find it useful to reference just the variable part of the density (integrating constant excluded) in which case we use the symbol ' $\propto$ ' to denote 'is proportional to.' In view of the prior-to-posterior conjugacy shared by many of the models we consider, we adopt the notational convention employed by Drèze and Richard (1983) wherein postscripts indicated ' $^o$ ' reflect prior information and postscripts indicated ' $^*$ ' reflect posterior information; accordingly  $f(\theta|\theta^o) \equiv \pi(\theta)$  and  $f(\theta|\theta^*) \equiv \pi(\theta|y)$ . Finally, we use indices  $i = 1, 2, \dots, N$ , to reference the households in question, where, we remind the reader,  $N = 401$ .

## A.2 Probability density functions

We use twelve probability density functions. The first, which we use to model the prior and posterior information about the single unknown *parameter* in the conjugate *Poisson sampling model* is the *Gamma distribution* (Bernardo and Smith, p. 118), namely,  $f_{1 \times 1}^{Ga}(x|\alpha, \beta) \equiv \beta^\alpha \Gamma(\alpha)^{-1} \lambda^{\alpha-1} \exp\{-\beta x\}$ ,  $\alpha \in (0, +\infty)$ ,  $\beta \in (0, +\infty)$ ,  $x \in (0, +\infty)$ . The second, which we use to model the data in the Poisson sampling model is the Poisson distribution (Bernardo and Smith, p. 116), namely,  $f_{1 \times 1}^{Pn}(x|\lambda) \equiv \exp\{-\lambda\} \lambda^x (x!)^{-1}$ ,  $\lambda \in (0, +\infty)$ ,  $x \in [0, 1, 2, \dots, +\infty)$ . The third, which we use to model the evidence in the Poisson sampling model is the so-called Poisson-Gamma distribution (Bernardo and Smith, p. 119),  $f_{1 \times 1}^{Pg}(x|\alpha, \beta, n) \equiv \beta^\alpha \Gamma(\alpha)^{-1} \Gamma(\alpha+x) (x!)^{-1} n^x (\beta+n)^{-(\alpha+x)}$ ,  $\alpha \in (0, +\infty)$ ,  $\beta \in (0, +\infty)$ ,  $x \in [0, 1, 2, \dots, +\infty)$ . The fourth, which we use to model the prior and posterior information about the single unknown parameter in the conjugate Geometric sampling model is the *Beta distribution* (Bernardo and Smith, p. 116), namely,  $f_{1 \times 1}^{Be}(x|\alpha, \beta) \equiv \Gamma(\alpha+\beta) \Gamma(\alpha)^{-1} \Gamma(\beta)^{-1} x^{\alpha-1} x^{\beta-1}$ ,  $\alpha \in (0, +\infty)$ ,  $\beta \in (0, +\infty)$ ,  $x \in (0, 1)$ . The fifth, which we use to model the data in the Geometric sampling model is the *Geometric distribution* (Bernardo and Smith, p. 116)  $f_{1 \times 1}^{Ge}(x|\theta) \equiv \theta (1-\theta)^x$ ,  $\theta \in (0, 1)$ ,  $x \in [0, 1, 2, \dots, +\infty)$ . The sixth, which we use to model the evidence in the Geometric sampling model is the *Geometric-Beta distribution* (Bernardo and Smith, p. 118),  $f_{1 \times 1}^{Gb}(x|\alpha, \beta, n) \equiv \Gamma(\alpha+\beta) \Gamma(\alpha)^{-1} \Gamma(\beta)^{-1}$

$\Gamma(\alpha+1) \Gamma(\beta+x) \Gamma(\alpha+\beta+x+1)^{-1}$ ,  $\alpha \in (0,+\infty)$ ,  $\beta \in (0,+\infty)$ ,  $x \in [0, 1, 2, \dots, +\infty)$ . The seventh, which we use to model the data in the Binomial sampling model, is the *Binomial distribution* (Bernardo and Smith, p. 115),  $f_{1 \times 1}^{\text{Bi}}(x|\theta, n) \equiv (n!) (x!)^{-1} ((n-x)!)^{-1} \theta^x (1-\theta)^{n-x}$ ,  $\theta \in (0,1)$ ,  $n \in [1, 2, \dots, +\infty)$ ,  $x \in [0, 1, 2, \dots, +\infty)$ . The eighth, which we use to model the evidence in the conjugate Binomial sampling model is the *Binomial-Beta distribution* (Bernardo and Smith, p. 117),  $f_{1 \times 1}^{\text{Bb}}(x|\alpha, \beta, n) \equiv \Gamma(\alpha+\beta) \Gamma(\alpha)^{-1} \Gamma(\beta)^{-1} \Gamma(\alpha+x) \Gamma(\beta+n-x) \Gamma(\alpha+\beta+n)^{-1}$ ,  $\alpha \in (0,+\infty)$ ,  $\beta \in (0,+\infty)$ ,  $x \in [0, 1, 2, \dots, +\infty)$ . The ninth, which we use to model the data in the Negative-Binomial sampling model is the *Negative-Binomial distribution* (Bernardo and Smith, p. 116),  $f_{1 \times 1}^{\text{Nb}}(x|\theta, r) \equiv ((r+x-1)!) (x!)^{-1} ((r-1)!)^{-1} \theta^r (1-\theta)^x$ ,  $\theta \in (0,1)$ ,  $r \in [1, 2, \dots, +\infty)$ ,  $x \in [0, 1, 2, \dots, +\infty)$ . The tenth, which we use to model the evidence in the conjugate Negative-Binomial sampling model is the *Negative-Binomial-Beta distribution* (Bernardo and Smith, p. 118),  $f_{1 \times 1}^{\text{Nbb}}(x|\alpha, \beta, r) \equiv \Gamma(\alpha+\beta) \Gamma(\alpha)^{-1} \Gamma(\beta)^{-1} \Gamma(\alpha+r) \Gamma(\beta+x) \Gamma(\alpha+\beta+x+r)^{-1}$ ,  $\alpha \in (0,+\infty)$ ,  $\beta \in (0,+\infty)$ ,  $r \in [1, 2, \dots, +\infty)$ ,  $x \in [0, 1, 2, \dots, +\infty)$ . The eleventh, which we use to model the prior information for the regression coefficients in the models containing covariates, is the *Multivariate-Normal distribution* (Bernardo and Smith, p. 136),  $f_{K \times 1}^{\text{Mn}}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \equiv (2\pi)^{-.5K} |\boldsymbol{\Sigma}|^{-.5} \exp\{ -.5(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \}$ . Finally, for the purpose of drawing from cumulative distribution functions and the application of the ‘probability integral transform method,’ we make use of the *Uniform distribution*  $f_{1 \times 1}^{\text{U}}(x|\alpha, \beta) \equiv (\beta-\alpha)^{-1}$ ,  $\alpha \in (-\infty, \beta)$ ,  $\beta \in (\alpha, +\infty)$ , which is the twelfth pdf we employ.

### A.3 Conjugate Poisson model development

The assumption that the data,  $y_1, y_2, \dots, y_N$ , follow the Poisson sampling distribution, leads to  $y_i \sim f_{1 \times 1}^{\text{Pn}}(y_i|\lambda)$ . The Poisson model has a conjugate prior probability density function, which is of the Gamma form. Hence, by assuming that  $\lambda \sim f_{1 \times 1}^{\text{Ga}}(\lambda|\alpha_o, \beta_o)$  and that the observations,  $y_1, y_2, \dots, y_N$ , are independent and identically distributed as exchangeable random variables, it follows that the posterior distribution for the Poisson parameter,  $\lambda$ , is also in the Gamma form,  $\lambda|\mathbf{y} \sim f_{1 \times 1}^{\text{Ga}}(\lambda|\alpha_*, \beta_*)$ , wherein  $\alpha_* = \alpha_o + \sum_{i=1}^N y_i$  and  $\beta_* = \beta_o + N$  define the location and scale of the posterior

distribution for  $\lambda$  which has mean and variance, namely  $E\{\lambda\} = \alpha_* \div \beta_*$  and  $\text{Var}\{\lambda\} = \alpha_* \div \beta_*^2$ .

The Poisson sampling model's marginal density for the data,  $f(\mathbf{y})$ , is in the Poisson-Gamma form,

$$f_{1 \times 1}^{\text{Pg}}(\mathbf{x}|\alpha, \beta, n) \equiv \beta^\alpha \Gamma(\alpha)^{-1} \Gamma(\alpha+x) (x!)^{-1} n^x (\beta+n)^{-(\alpha+x)}, \alpha \in (0, +\infty), \beta \in (0, +\infty), x \in [0, 1, 2, \dots, +\infty).$$

Finally, in the Gamma prior pdf,  $\lambda \sim f_{1 \times 1}^{\text{Ga}}(\lambda|\alpha_o, \beta_o)$ , as in subsequent prior specifications, we set  $\alpha_o = \beta_o = 1$ .

#### ***A.4 Conjugate Geometric model development***

Developments for the conjugate Geometric-Beta sampling model follow analogously with that of the Poisson model. The data,  $y_1, y_2, y_N$ , are assumed to follow the Geometric distribution,

$$f_{1 \times 1}^{\text{Ge}}(y_i|\theta) \equiv \theta (1-\theta)^{y_i}, \theta \in (0,1), y_i \in [0, 1, 2, \dots, +\infty). \text{ The unknown parameter in the Geometric}$$

distribution, which is  $\theta$ , is assumed to follow the prior probability density which is the Beta

distribution, namely,  $f_{1 \times 1}^{\text{Be}}(\theta|\alpha_o, \beta_o)$ ,  $\alpha_o \in (0, +\infty)$ ,  $\beta_o \in (0, +\infty)$ ,  $\theta \in (0,1)$ . The conjugate posterior

is in the Beta form and is  $f_{1 \times 1}^{\text{Be}}(\theta|\alpha_*, \beta_*)$ , wherein  $\alpha_* = \alpha_o + N$  and  $\beta_* = \beta_o + \sum_{i=1}^N y_i$  define the

location and scale of the posterior distribution for  $\theta$  for which the mean and variance are,

respectively,  $E\{\theta\} = \alpha_* \div (\alpha_* + \beta_*)$  and  $\text{Var}\{\theta\} = \alpha_* \beta_* \div (\alpha_* + \beta_* + 1) \div (\alpha_* + \beta_*)^2$ . The

Geometric sampling model has a marginal likelihood,  $f(\mathbf{y})$ , in the Negative-Binomial-Beta form,

$$f_{1 \times 1}^{\text{NbB}}(\mathbf{y}|\alpha_o, \beta_o, N, r), \text{ which is available by direct calculation. Finally, in the Beta prior pdf, } \theta \sim$$

$$f_{1 \times 1}^{\text{be}}(\lambda|\alpha_o, \beta_o), \text{ we set } \alpha_o = \beta_o = 1.$$

#### ***A.5 Conjugate Negative Binomial model development***

Developments concerning the Negative Binomial formulation follow analogously to the Geometric

model when it is recognized that the Geometric distribution is a special case of the Negative

Binomial formulation with the restriction  $r_o = 1$ . Analyses follow directly by extension to those

described under the Geometric setting with the restriction  $r_o = 1$  relaxed to  $r_o \neq 1$ . Once again, the

conjugate Negative Binomial model has a marginal likelihood which is available in closed form,

thus facilitating straight-forward models comparisons.

### A.6 Conjugate Binomial model development

Developments concerning the Binomial formulation are as follows. The data,  $y_1, y_2, y_N$ , are assumed to be distributed as Binomial random deviates,  $f_{1 \times 1}^{\text{Bi}}(x|\theta, n) \equiv (n!) (x!)^{-1} ((n-x)!)^{-1} \theta^x (1-\theta)^{n-x}$ ,  $\theta \in (0,1)$ ,  $n \in [1, 2, \dots, +\infty)$ ,  $x \in [0, 1, 2, \dots, +\infty)$ . The unknown parameter in the Binomial distribution,  $\theta$ , is assumed to follow the prior probability density which is the Beta distribution, namely,  $f_{1 \times 1}^{\text{Be}}(\theta|\alpha_o, \beta_o)$ ,  $\alpha_o \in (0, +\infty)$ ,  $\beta_o \in (0, +\infty)$ ,  $\theta \in (0,1)$ . The conjugate posterior is in the Beta form  $f_{1 \times 1}^{\text{Be}}(\theta|\alpha_*, \beta_*)$ ,  $\alpha_* \in (0, +\infty)$ ,  $\beta_* \in (0, +\infty)$ ,  $\theta \in (0,1)$ , wherein  $\alpha_* = \alpha_o + N$  and  $\beta_* = \beta_o + \sum_{i=1}^N y_i$  and the marginal likelihood is in the Binomial-Beta form  $f_{1 \times 1}^{\text{Bb}}(x|\alpha, \beta, n) \equiv \Gamma(\alpha+\beta) \Gamma(\alpha)^{-1} \Gamma(\beta)^{-1} \Gamma(\alpha+x) \Gamma(\beta+n-x) \Gamma(\alpha+\beta+n)^{-1}$ ,  $\alpha \in (0, +\infty)$ ,  $\beta \in (0, +\infty)$ ,  $x \in [0, 1, 2, \dots, +\infty)$ . Finally, in the Beta prior pdf,  $\theta \sim f_{1 \times 1}^{\text{be}}(\lambda|\alpha_o, \beta_o)$ , we set  $\alpha_o = \beta_o = 1$ .

### A.7 Generalized linear models development

In the Poisson setting, inferences about the all-important location-and-scale parameter,  $\lambda$ , are attainable by making three very small modifications to the conjugate Poisson-Gamma sample setting. First, we allow  $\lambda$  to vary across respondents, assuming that there exist quantities,  $\lambda_1, \lambda_2, \dots, \lambda_N$  depicting the propensity for each household to ‘adopt.’ Second, we assume, in turn, that each choice is conditioned by a K-vector of household-specific covariates,  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ , consisting of household specific elements,  $\mathbf{x}_i \equiv (x_{i1}, x_{i2}, \dots, x_{iK})'$ . Third, we consider that the adoption propensities, which are unobserved, are affected by a vector of unobserved coefficients  $\boldsymbol{\beta} \equiv (\beta_1, \beta_2, \dots, \beta_K)'$  which are assumed to evolve, in turn, from a K-variate Normal prior distribution  $\boldsymbol{\beta} \sim f_{K \times 1}^{\text{Mn}}(\boldsymbol{\beta}|\boldsymbol{\mu}_o, \boldsymbol{\Sigma}_o)$ , wherein  $\boldsymbol{\mu}_o$  denotes the K-dimensional null-vector and  $\boldsymbol{\Sigma}_o$  denotes the K-dimensional identity matrix scaled by a factor of 10. Finally, observing that the  $\lambda_1, \lambda_2, \dots, \lambda_N$  must satisfy  $\lambda_1 > 0, \lambda_2 > 0, \dots, \lambda_N > 0$ , we assume that for each  $i = 1, 2, \dots, N$ ,  $\lambda_i \equiv \exp\{\mathbf{x}_i' \boldsymbol{\beta}\}$ . The model is no longer conjugate and so inferences, including conclusions about the evidence in favor of any particular model, must be derived by iterative simulation, such as Markov Chain Monte Carlo. Notwithstanding this remark,

computational issues are rather modest in this generalized-linear setting. It follows from work that is now considered quite standard in such situations, (Casella and Robert, 1999; Koop, 2003; Koop, Poirier and Tobias, 2007) that estimates of the full model, with joint distribution  $f(\mathbf{y}, \boldsymbol{\beta} | \mathbf{x}) = \prod_{i=1}^N f_{1 \times 1}^{\text{Pn}}(y_i | \exp\{\mathbf{x}_i' \boldsymbol{\beta}\}) \times f_{K \times 1}^{\text{Mn}}(\boldsymbol{\beta} | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$  are easily retrieved. In particular, a basic Metropolis-Hastings algorithm for retrieving estimates of  $\boldsymbol{\beta}$ ; for providing a foundation for models comparison extensions (Jeliazkov and Chib, 2001); and for subsequent posterior prediction and inference; proceeds sequentially, for iterations  $t = 0, t=1, t=2, \dots$ , as follows:  $\langle\langle \text{Algorithm}(A_o) \rangle\rangle \langle\langle \text{Given the state variable at iteration 't', say } \boldsymbol{\beta}^{(t)}; \text{ draw a K-variate proposal, say } \boldsymbol{\delta}^{(t)}, \text{ from the K-variate-Normal distribution, } f_{K \times 1}^{\text{Mn}}(\boldsymbol{\delta}^{(t)} | \boldsymbol{\beta}^{(t)}, \mathbf{I}_K \sigma^{(t)}), \text{ where } \sigma^{(t)} \text{ denotes an investigator-supplied variance parameter. Compute the ratio } \mathfrak{R}^{(t)} \equiv f(\mathbf{y}, \boldsymbol{\delta}^{(t)} | \mathbf{x}) \div f(\mathbf{y}, \boldsymbol{\beta}^{(t)} | \mathbf{x}). \text{ Accept the draw such that } \boldsymbol{\beta}^{(t+1)} = \boldsymbol{\delta}^{(t)} \text{ with a pre-specified probability, say, } \wp^{(t)}; \text{ and set } \boldsymbol{\beta}^{(t+1)} = \boldsymbol{\delta}^{(t)}, \text{ with probability } 1 - \wp^{(t)}. \rangle\rangle$  This algorithm is very simple to implement and leads to robust estimates of the regression coefficients in the generalized-linear Poisson setup. In practice, we target the acceptance rate in the proposal distribution to be fixed at  $\wp^{(t)} = \wp = 0.25$  and we allow adjustments in the variance terms  $\sigma^{(t)}, \sigma^{(t+1)}, \sigma^{(t+2)}$ , such that this acceptance rate is targeted throughout the simulation. In the specific case of the Poisson distribution, we find that estimates are rather quick to converge, requiring only modest sample sizes slightly in excess of about 1,000 iterations. However, in all of the reports enacted with respect of the algorithm we sample conservatively for an initial 10,000 iterations, using the subsequent 10,000 iterations for reporting. We note that, one attractive feature of this algorithm, due to Chib and Jeliazkov (2001) is that it is relatively straight-forward to derive robust estimates of the marginal likelihood attributable to a specific model by simply extending the sampling scheme for a subsequent two rounds of draws. In summary, closed-form analysis of the Poisson-Gamma sampling framework facilitates inference, posterior prediction and model selection. However, the conjugate Poisson-Gamma model is restrictive. Importantly, it precludes the evaluation of potentially important and fundamental covariate information. A generalized-linear Poisson-Normal model is available from minor modifications of the Poisson-Gamma configuration and the resulting

joint posterior distribution is easily assessed using routine application of Markov Chain Monte Carlo techniques. The generalized-linear Poisson-Normal model leads to improvements in fit. Implementation under Geometric, Negative Binomial and Binomial sampling schemes is virtually identical, except for one major modification. Recognizing that the sample space for these sampling models is  $\theta \in (0,1)$ , we implement the probability measure using a logit transformation of the regression means. That is,  $\theta_1, \theta_2, \dots, \theta_N$ , are modeled using the conditions  $\theta_i \equiv \exp\{\sum_j x_{ji}\beta_j\} \div \{1+\exp\{\sum_j x_{ji}\beta_j\}\}$ ,  $i = 1, 2, \dots, N$ . And all of the remaining detail outlined in the context of the Poisson generalized linear model proceeds analogously. The computer codes used to implement the relevant simulations are available within the supplemental materials accompanying this publication. Additional inferential detail is available from the authors upon request.

## References

- Abdulai, A., Huffman, W. E., 2005. The diffusion of new agricultural technologies: The case of crossbred-cow technology in Tanzania, *Amer. J. Agr. Econ.* 87, 645-659.
- Ahmed, M., Ehui, S., Saleem, M., 2002. Adoption of crossbred cow technologies and increased food security among smallholder dairy farmers in East African highlands, *J Food Prod.* 6, 319-337.
- Albert, J. and S. Chib. 1993. Bayesian analysis of dichotomous and polychotomous response data. *Journal of the American Statistical Association*, 88, 422, 669-679.
- Barrett, C.B. 1996. On price risk and the inverse farm size-productivity relationship. *Journal of Development Economics*, 51 (2), 193-215
- Bernardo, J., 1996. The concept of exchangeability and its applications. *Far East J. Mathematical Sciences* 4, 111-121.
- Besley, T., Case, A., 1993. Modeling technology adoption in developing countries, *Amer. Econ. Rev.* 83, 396.
- Chib, S., 1995. Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association* 90(432):1313-1321.

- Chib, S., Jeliazkov, I., 2001. Marginal likelihood from the Metropolis-Hastings output. *Journal of the American Statistical Association*, 96, 270-81.
- de Finetti, B. , 1938. ‘Sur la Condition d’équivalence Partielle’. *Actualités Scientifiques et Industrielles* 739. Paris: Herman and Cii.
- Department.of.Census.and.Statistics, 2002. *Census of Agriculture 2002: Smallholding Sector*. Department of Census and Statistics, Sri Lanka, Colombo.
- Dinar, A. and G. Keynan. 2001. Economics and paid extension: Lessons from experience in Nicaragua. *American Journal of Agricultural Economics*, 83, 3, 769-776.
- Doss, C. R., 2006. Analyzing technology adoption using microstudies: limitations, challenges, and opportunities for improvement, *Agr. Econ.* 34, 207-219.
- Drèze, J. H., Richard, J.-F., 1983. Bayesian Analysis of Simultaneous Equations Systems in *Handbook of Econometrics*, in: Z. Griliches & M. D. Intriligator (eds.), *Handbook of Econometrics*, edition 1, volume 1, chapter 9, pages 517-598. Amsterdam: Elsevier.
- Edirisinghe, J. C. 2010. Technology adoption, efficiency and market access of smallholder dairy producers in Sri Lanka: A Bayesian perspective. PhD dissertation submitted in partial fulfilment of the requirements for the doctoral degree, School of Agriculture, Policy and Development, University of Reading, United Kingdom.
- Ehui, S., H. Li-Pun, V. Mares, and B. Shapiro. 1998. The role of livestock in food security and environmental protection. *Outlook on Agriculture*, 27(2): 81-87
- Feder, G., Just, R. E., Zilberman, D., 1985. Adoption of agricultural innovations in developing countries: A survey, economic development & cultural change. 33, 255.
- Feder, G., Slade, R., 1984. The acquisition of information and the adoption of new technology, *Amer. J. Agr. Econ.* 66, 312.
- Foltz, J. D., Chang, H., 2002. The adoption and profitability of rbST on Connecticut dairy farms, *Amer. J. Agr. Econ.* 84, 1021-1032.

- Frisvold, G., K. Fericola and M. Langworthy. 2001. Market returns, infrastructure and the supply and demand for extensions services. *American Journal of Agricultural Economics*, 83, 3, 758-63.
- Giunta, A., Trivieri, F., 2007. Understanding the determinants of information technology adoption: evidence from Italian manufacturing firms, *Appl. Econ.* 39, 1325 - 1334.
- Hanson, J. and R. Just. 2001. The potential for transition to paid extension: Some guiding economic principles. *American Journal of Agricultural Economics*, 83, 3, 777-784.
- Hattam, C. E., D. J. Lacombe and G. J. Holloway, 2012. Organic certification, export-market access and the impacts of policy: Bayesian estimation of avocado smallholder “times-to-organic certification” in Michoacán Mexico. *Agricultural Economics*, 43, 441-457.
- Holloway, G. J., Barrett, C. B., Ehui, S., Kenya, A., 2000a. Cross-bred cow adoption and milk-market participation in a multivariate, count-data framework, *Bayesian Methods with Applications to Science, Policy, and Official Statistics, Selected Papers from ISBA, Eurostat*, 23-241.
- Holloway, G., J., Dorfman, J. H., Ehui, S. K., 2001. Tobit estimation with unknown point of censoring with an application to milk market participation in the Ethiopian highlands, *Can. J. Agr. Econ.* 49, 293-311.
- Holloway, G. J. and S. K. Ehui. 2001. Demand, supply and willingness-to-pay for extension services in an emerging-market setting. *American Journal of Agricultural Economics* 83, 3, 764-68.
- Holloway, G. and S. Ehui. 2002. Expanding market participation among smallholder livestock producers: A collection of studies employing Gibbs sampling and data from the Ethiopian highlands (1998-2001). Addis Ababa: International Livestock Research Institute Working Paper Series, 2002.



- Holloway, G., Ehui, S. and Teklu, A., 2008. Bayes estimates of distance to market: Transactions costs, cooperatives and milk-market development in the Ethiopian highlands, *Journal of Applied Econometrics* 23, 683-696.
- Holloway, G. J., C. Nicholson, C. Delgado, S. Ehui and S. Staal. 2000. Agroindustrialization through institutional innovation: Transactions costs, cooperatives and milk-market development in the East African highlands. *Agricultural Economics* 23, 279-88.
- Johnson, W. E. 1924. *Logic: Part III. The Logical Foundations of Science*. Cambridge: Cambridge University Press.
- Kiwuwa G H, Trail J C M, Kurtu M Y, Worku G, Anderson F M, Durkin J. 1983. Cross-Breed Dairy Cattle Productivity in Arsi Region, Ethiopia. International Livestock Center for Africa (ILCA) Research Report No. 11. Addis Ababa: ILCA.
- Koop, G., 2003. *Bayesian Econometrics*. West Sussex. John Wiley and Sons Ltd.
- Koop, G., D. Poirier and J. Tobias. 2007. *Bayesian Econometric Methods*. Cambridge: Cambridge University Press.
- Koundouri, P., Nauges, C. I., Tzouvelekas, V., 2006. Technology adoption under production uncertainty: Theory and application to irrigation technology, *Amer. J. Agr. Econ.* 88, 657-670.
- Kreps, D. M. 1988. *Notes on The Theory of Choice*. Underground Classics in Economics. Boulder, Colorado: Westview Press.
- Lindley, D. and A. F. M. Smith. 1972. Bayes estimates for the linear model. *Journal of the Royal Statistical Society*, 34, 1, 1-41.
- Lal, K., 1999. Determinants of the adoption of information technology: a case study of electrical and electronic goods manufacturing firms in India, *Research Policy*. 28, 667-680.
- Lapar, M. L. A., Ehui, S. K., 2004. Factors affecting adoption of dual-purpose forages in the Philippine uplands, *Agr Syst.* 81, 95-114.
- Misra, S. K., Carley, D. H., Fletcher, S. M., 1993. Factors influencing southern dairy farmers' choice of milk handlers, *Journal of Agricultural and Applied Economics*. 25, 197-197.

- Mood, A. M., F. A. Graybill and D. C. Boes. 1974. Introduction to the theory of Statistics. New York; McGraw-Hill.
- Negatu, W., Parikh, A., 1999. The impact of perception and other factors on the adoption of agricultural technology in the Moret and Jiru Woreda (district) of Ethiopia, *Agr. Econ.* 21, 205-216.
- Nicholson C F. 1997. The Impact of Milk Groups in the Shewa and Arsi Regions of Ethiopia: Project Description, Survey Methodology, and Collection Procedures. Mimeograph, Livestock Policy Analysis Project, International Livestock Research Institute, Addis Ababa.
- Nicholson, C. F., Thornton, P. K., Muinga, R. W., 2004. Household-level impacts of dairy cow ownership in coastal Kenya, *J. Agr. Econ.* 55, 175-195.
- Poirier, D. J., 1988. Frequentist and subjectivist perspectives on the problems of model building in Economics (with discussion). *Journal of Economic Perspectives*, 2: 121-170.
- Poirier, D. J., 1995. Intermediate Statistics and Econometrics: A Comparative Approach. Cambridge: MIT Press.
- Poirier, D. J., 2011. Exchangeability, representation theorems and subjectivity, in Geweke, J., Koop, G. and van Dijk, H. (eds.) The Oxford Handbook of Bayesian Econometrics, Oxford; Oxford University Press, 2011.
- Singh, I., L. Squire and J. Strauss. Agricultural Household Models. Washington, DC: World Bank, 1986.
- Staal, S. J., Baltenweck, I., Waithaka, M. M., deWolff, T., Njoroge, L., 2002. Location and uptake: integrated household and GIS analysis of technology adoption and land use, with application to smallholder dairy farms in Kenya, *Agr. Econ.* 27, 295-315.
- Sunding, D., Zilberman, D., 2001. The agricultural innovation process: research and technology adoption in a changing agricultural sector, in B. L. Gardner and G. C. Rausser (eds.) *The Handbook of Agricultural Economics, Handbooks In Economics. Volume 18*, 207-262. Amsterdam: North-Holland, 2001.

Zellner A. 1996. An Introduction to Bayesian Inference in Econometrics. Wiley Classics Library.  
New York: John Wiley and Sons.

## Tables

Table 1: Description of variables

Variable (number)	Description
Cow number	Total number of cows owned by the household
Dry Zone (1)	Household live in the dry zone = 1
Wet Zone (2)	Household live in the wet zone = 1
Intermediate Zone (3)	Household live in the intermediate zone = 1
Time to market (4)	Time to the closest buyer in minutes
Total Income (5)	Code for the Total income of the household in Sri Lankan Rupees per month 1. below 5000; 2. 5000 – 10000; 3. 11000-15000; 4. 16000- 20000; 5. 21000 – 25000; 6. 26000- 30000; 7. 31000 – 35000; 8. above 35000
Land (6)	Land extent owned
Cattle Shed (7)	Owns a cattle shed =1
Perception I (8)	Perception of the farmer on the statement : “Cross bred animals are not suitable in your area” 1=do not agree ; 2=don’t know; 3=agree
Perception II (9)	Perception of the farmer on the statement : “It is expensive to maintain cross bred animals”
Society (10)	Household is a member of milk society = 1
Training (11)	Household had training on dairy activities = 1
Education (12)	Number of years in formal education
Experience (13)	No of years farmer has been engaged in dairy farming

Table 2: Summary statistics stratified by agro-ecological zones

	Dry Zone		Wet Zone		Intermediate Zone		All Data	
	N = 152		N = 93		N = 156		N = 401	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Cow number	1.61	1.47	2.03	1.59	1.76	1.61	1.65	1.54
Dry zone	1.00	0.00	0.00	0.00	0.00	0.00	0.29	0.45
Wet Zone	0.00	0.00	1.00	0.00	0.00	0.00	0.30	0.46
Intermediate Zone	0.00	0.00	0.00	0.00	1.00	0.00	0.41	0.49
Time to market	11.09	11.57	13.93	14.37	22.01	37.07	21.78	36.75
Total Income	3.82	2.05	3.28	1.85	3.66	2.01	3.60	2.02
Land	2.14	2.18	1.45	1.81	1.45	1.65	1.56	1.68
Cattle Shed	0.59	0.49	0.81	0.40	0.78	0.42	0.64	0.48
Perception I	1.42	0.80	1.28	0.68	1.40	0.80	1.39	0.77
Perception II	2.87	0.47	2.46	0.89	2.78	0.63	2.60	0.79
Society	0.76	0.43	0.87	0.34	0.69	0.46	0.74	0.44
Training	0.66	0.47	0.66	0.48	0.59	0.49	0.65	0.48
Education	7.53	3.01	8.18	2.75	7.80	2.88	7.49	3.00
Experience	12.22	10.72	15.72	9.25	13.99	11.09	14.37	10.45

Note: Std. = Standard Deviation

Table 3: Summary statistics stratified by adoption class

	Non Adopters		Adopters		All Sample	
	N = 51		N = 350		N = 401	
	Mean	Std.	Mean	Std.	Mean	Std.
Cow number	0.00	0.00	1.00	0.00	1.71	1.53
Dry zone	0.51	0.50	0.35	0.48	0.27	0.44
Wet Zone	0.10	0.30	0.25	0.43	0.34	0.47
Intermediate Zone	0.39	0.49	0.40	0.49	0.39	0.49
Time to market	41.08	53.75	12.95	14.97	20.77	34.90
Total Income	2.88	1.82	3.29	1.85	3.54	1.97
Land	1.69	1.77	1.40	1.54	1.54	1.63
Cattle Shed	0.31	0.46	0.77	0.42	0.66	0.47
Perception I	1.63	0.91	1.36	0.76	1.38	0.77
Perception II	2.67	0.70	2.79	0.60	2.58	0.80
Society	0.49	0.50	0.77	0.42	0.76	0.43
Training	0.35	0.48	0.57	0.50	0.43	0.65
Education	5.96	3.22	7.61	2.90	7.59	2.93
Experience	13.66	11.79	13.65	10.54	14.24	10.15

Note: Std. = Standard Deviation

Table 4. Generalized linear regression results.

	Specification														
	Poisson			Geometric			Negative Binomial			Binomial			Probit		
Dry Zone	-0.61	-0.30	0.03	-0.35	0.37	1.11	<b>1.72</b>	<b>1.96</b>	<b>2.53</b>	<b>-5.64</b>	<b>-5.24</b>	<b>-4.89</b>	-1.66	-0.62	0.41
Wet Zone	-0.28	-0.01	0.31	-0.61	0.14	0.86	<b>1.37</b>	<b>1.71</b>	<b>2.17</b>	<b>-5.47</b>	<b>-4.97</b>	<b>-4.53</b>	-0.91	0.12	1.19
Intermediate Zone	-0.42	-0.14	0.19	-0.56	0.23	1.01	<b>1.49</b>	<b>1.84</b>	<b>2.43</b>	<b>-5.59</b>	<b>-5.09</b>	<b>-4.71</b>	-1.21	-0.15	0.91
Time to Market	<b>-2.87</b>	<b>-1.74</b>	<b>-0.84</b>	<b>0.33</b>	<b>1.41</b>	<b>2.43</b>	<b>0.33</b>	<b>1.55</b>	<b>2.30</b>	<b>-2.20</b>	<b>-1.49</b>	<b>-0.87</b>	<b>-4.04</b>	<b>-2.63</b>	<b>-1.32</b>
Total Income	<b>0.52</b>	<b>0.88</b>	<b>1.16</b>	<b>-1.44</b>	<b>-0.90</b>	<b>-0.30</b>	<b>-1.22</b>	<b>-0.81</b>	<b>-0.47</b>	<b>0.74</b>	<b>0.97</b>	<b>1.14</b>	-0.58	0.29	1.16
Land	<b>0.03</b>	<b>0.67</b>	<b>1.12</b>	-1.51	-0.48	0.66	-0.92	-0.42	0.10	-0.15	0.47	1.01	-2.24	-0.30	1.82
Cattle Shed	<b>0.00</b>	<b>0.15</b>	<b>0.31</b>	-0.48	-0.19	0.09	-0.29	-0.12	0.07	<b>0.00</b>	<b>0.17</b>	<b>0.36</b>	<b>0.29</b>	<b>0.71</b>	<b>1.13</b>
Perception I	-0.49	-0.24	0.06	-0.50	0.08	0.60	-0.06	0.21	0.49	-0.44	-0.19	0.15	-0.96	-0.29	0.40
Perception II	-0.49	-0.12	0.20	-0.43	0.15	0.63	<b>0.00</b>	<b>0.25</b>	<b>0.56</b>	-0.36	-0.14	0.07	-0.46	0.41	1.25
Society	-0.14	0.06	0.23	-0.43	-0.05	0.30	-0.25	-0.08	0.14	-0.14	0.06	0.26	-0.17	0.24	0.67
Training	<b>0.16</b>	<b>0.30</b>	<b>0.46</b>	-0.65	-0.33	0.00	<b>-0.52</b>	<b>-0.28</b>	<b>-0.09</b>	<b>0.14</b>	<b>0.31</b>	<b>0.46</b>	<b>0.07</b>	<b>0.49</b>	<b>0.92</b>
Education	-0.06	0.36	0.69	-1.11	-0.36	0.41	-0.73	-0.21	0.28	-0.13	0.26	0.73	<b>0.04</b>	<b>0.91</b>	<b>1.77</b>
Experience	-0.57	-0.07	0.34	-0.65	0.01	0.67	-0.68	-0.02	0.51	-0.28	0.12	0.69	-0.36	0.77	1.98
Count log marginal likelihood	-646.78			-731.75			-646.62			-643.33					
Numerical standard error	0.22			0.36			0.00			0.00					
Binary log marginal likelihood	-154.66			-100.48			-138.33			-150.92			-237.80		
Numerical standard error	0.00			0.00			0.04			0.03			4.05		

Note: Estimates derived from a Gibbs sample size of  $S = 10,000$  discarding an initial ‘burnin’ of  $S = 10,000$ . Within each specification, the columns report, respectively, the 2.5 % lower highest-posterior-density limit; the posterior mean; and the 97.5% upper highest-posterior-density limit. Emboldened entries are ‘significant’ at the 95% confidence level.’ ‘Significant’ within each specification are as follows: Poisson regression model: Time to Market, Total Income, Land, Cattle Shed and Training; Geometric regression model: Time to market, Total Income; Negative Binomial regression model: Dry Zone, Wet Zone, Intermediate Zone, Time to Market, Total Income, Perception II, Training. Binomial regression model: Dry Zone, Wet Zone, intermediate Zone, Time to Market, Total Income, Cattle Shed, Training; Probit regression model: Time to Market, Cattle Shed, Training, Education. The dominant count-data model is the Binomial regression model. The dominant binary-data model is the binary-transformed Geometric model.

## Figures

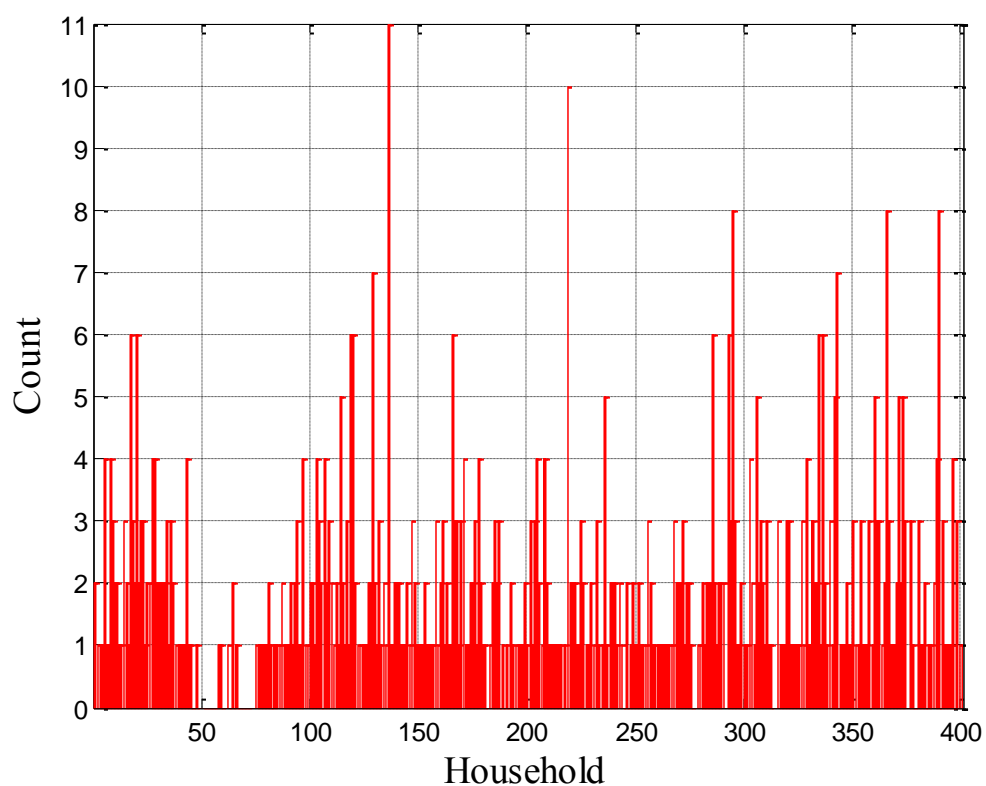


Figure 1. Distribution of crossbred cows across the sample. There are 401 households within the sample. The maximum number of cows is eleven; the minimum number of cows is zero; the mean number of cows is 1.77; the median is 1.00; and the standard deviation is 1.57.



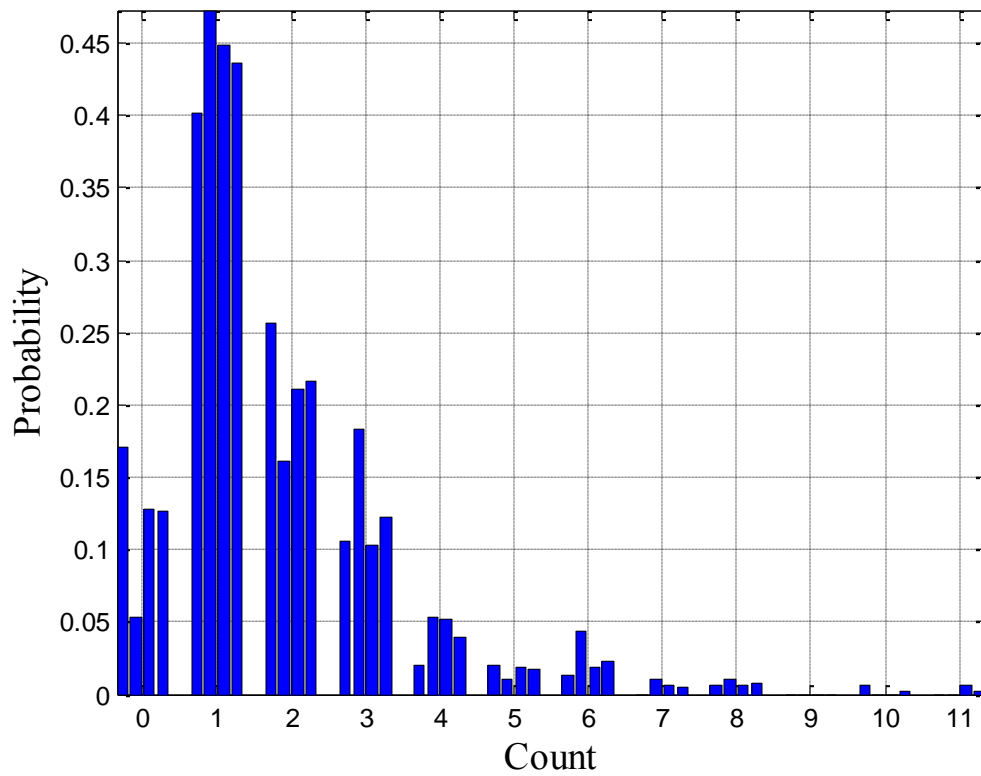


Figure 2. Distribution of crossbred cows across agro-climatic zones. The entries within each count are, respectively, the conditional probability of that count across the respective climatic zones and the full data set. The first bar denotes the probability of observing that count across the ‘dry zone,’ consisting of  $N = 152$  respondents; the second bar denotes the probability of observing that count within the ‘wet zone,’  $N = 93$ ; the third bar denotes the probability of observing that count within the ‘intermediate zone,’  $N = 156$ ; and the fourth bar denotes the empirical probability of observing that count across the overall sample,  $N = 401$ .

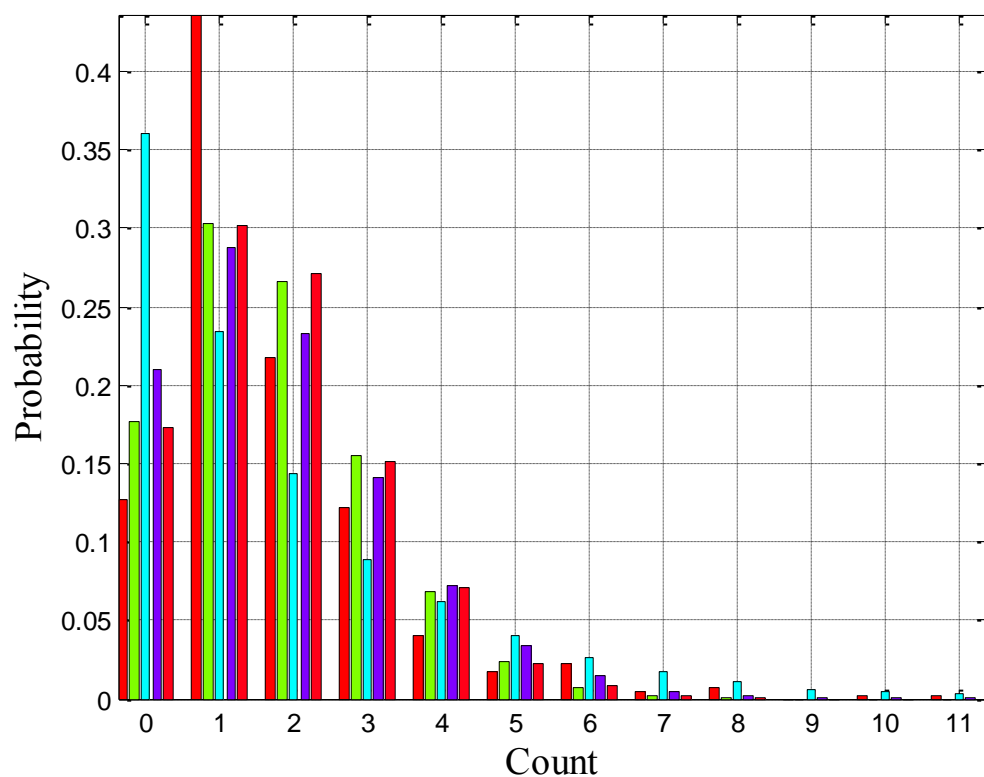


Figure 3. Crossbred-cow counts. The observed probabilities across the sample are reported for each count by the left-most (red) bar. The predicted probabilities derived from the Poisson model are reported in the (green) second bar. The predicted probabilities derived from the Geometric model are reported in the (cyan) third bar. The predicted probabilities derived from the Negative Binomial model are reported in the (magenta) fourth bar. The predicted probabilities derived from the Binomial model are reported in the right-most (red) bar.

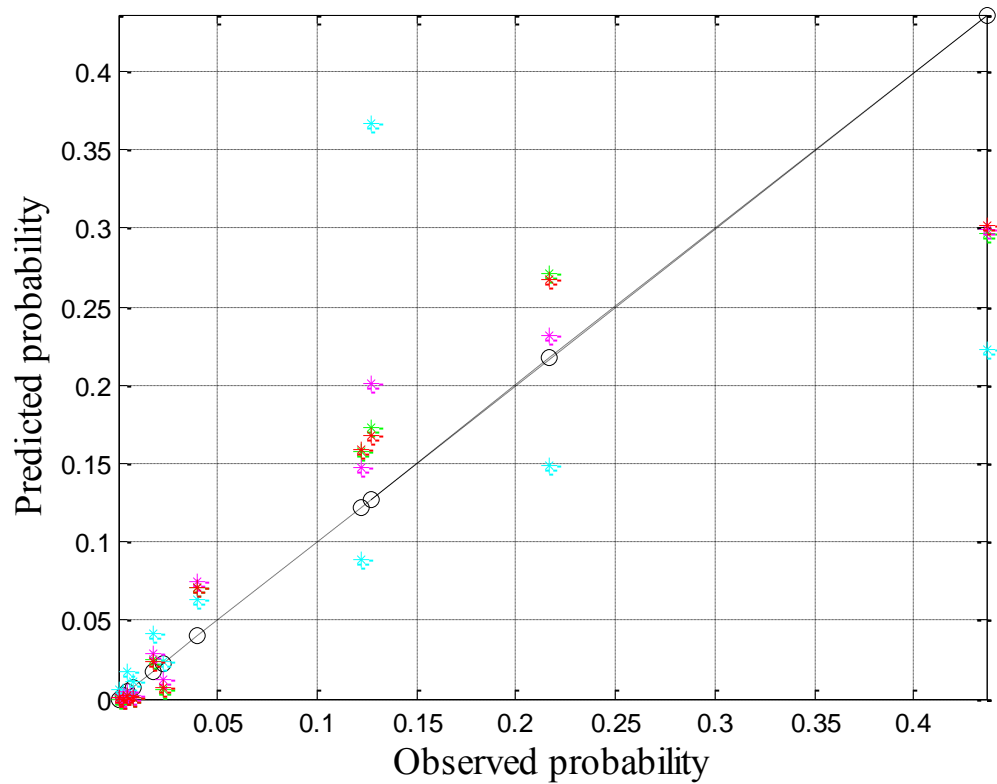


Figure 4. Empirically-observed and model-predicted probabilities. The black dotted line represents the line of ‘perfect fit’ with the empirically observed probabilities entered in circles. The predicted probabilities derived from the Poisson model are reported in green. The predicted probabilities derived from the Geometric model are reported in cyan. The predicted probabilities derived from the Negative Binomial model are reported in magenta. The predicted probabilities derived from the Binomial model are reported in red. The correlations between the empirically observed probabilities and the predicted probabilities for the Poisson, Geometric, Negative Binomial and Binomial models are, respectively, .927, .689, .926, and .929.

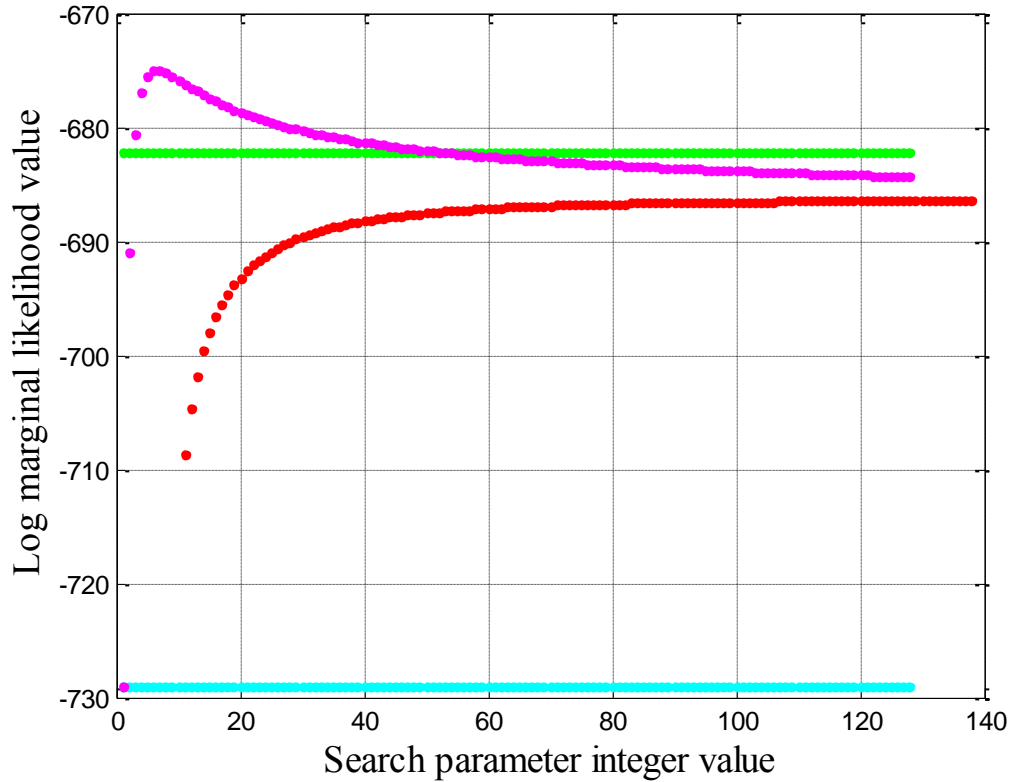


Figure 5. Log-marginal likelihood values for the conjugate Poisson, Geometric, Negative Binomial and Binomial models over iterated, integer-valued parameter space. The green entry depicts the constant log-marginal likelihood value corresponding to the Poisson model and the cyan entry depicts the constant log-marginal likelihood value corresponding to the Geometric model. The magenta entries depict the log-marginal likelihood value of the Negative-Binomial model as the integer-valued parameter,  $r$ , is varied from a lower bound of 1, at which point the Geometric and Negative Binomial models coincide; the maximum log-marginal likelihood value is attained at  $r = 7$ . The red entries correspond to the Binomial model, with the number of trials,  $n_o$ , adjusted uniformly across the sample from a lower bound of  $n_o = 11$ , which is the maximum number of crossbred animals held by a household; the maximum log-marginal likelihood value is attained at the maximum log-marginal likelihood value is attained at  $n_o = 139$ . Not indicated in the graphic is the fact that the Negative Binomial and Binomial log marginal likelihood values appear asymptotic. Conditional probabilities in favor of the Poisson, Geometric, Negative Binomial and Binomial models with the integer parameters set at their optimal values are 0.00078, 0.00000, 0.99921 and 0.00001.

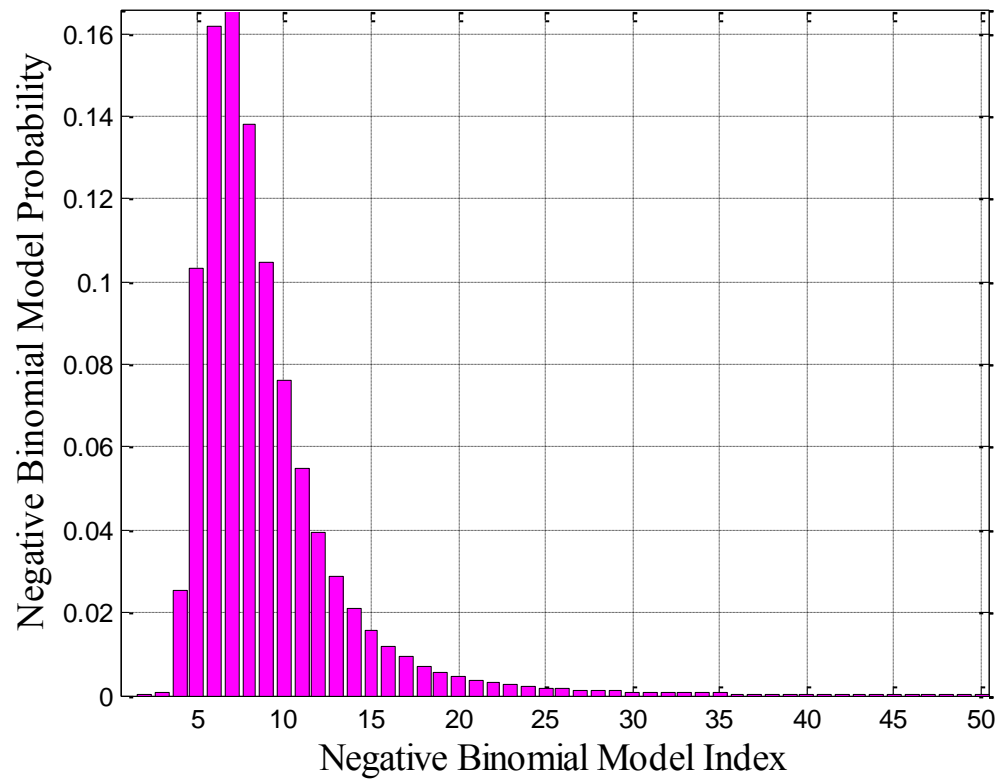


Figure 6. Model-average probabilities available from applying the integer-value search procedure documented in figure 4. The maximum probability over the first 50 iterations is 0.1654.

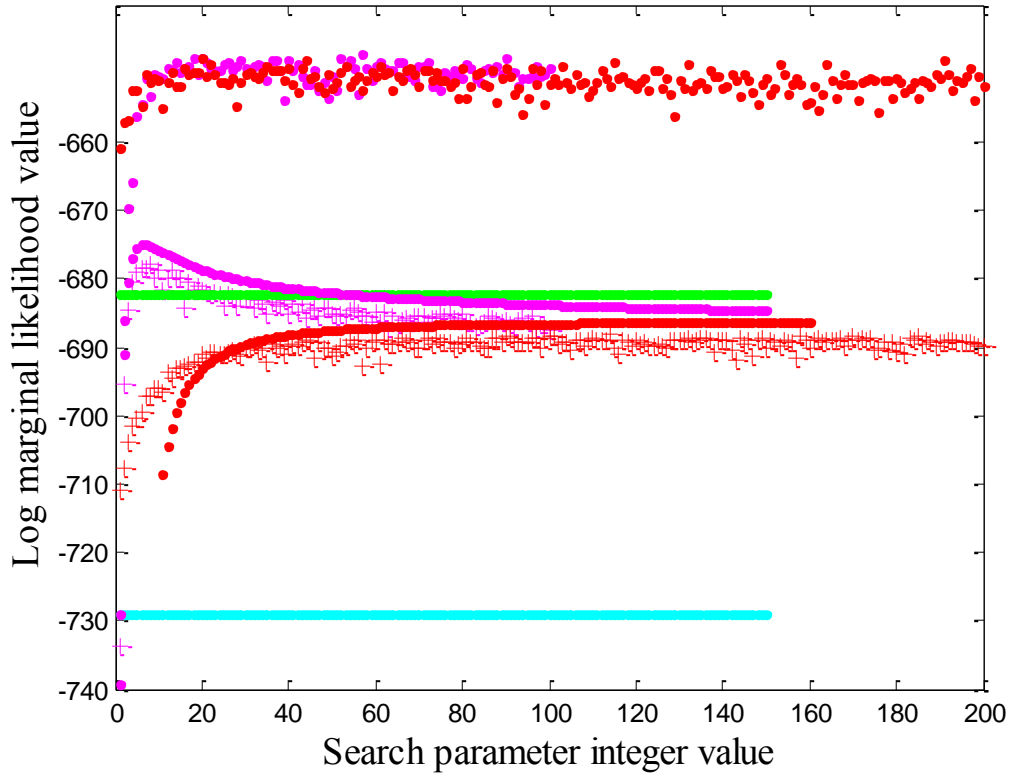


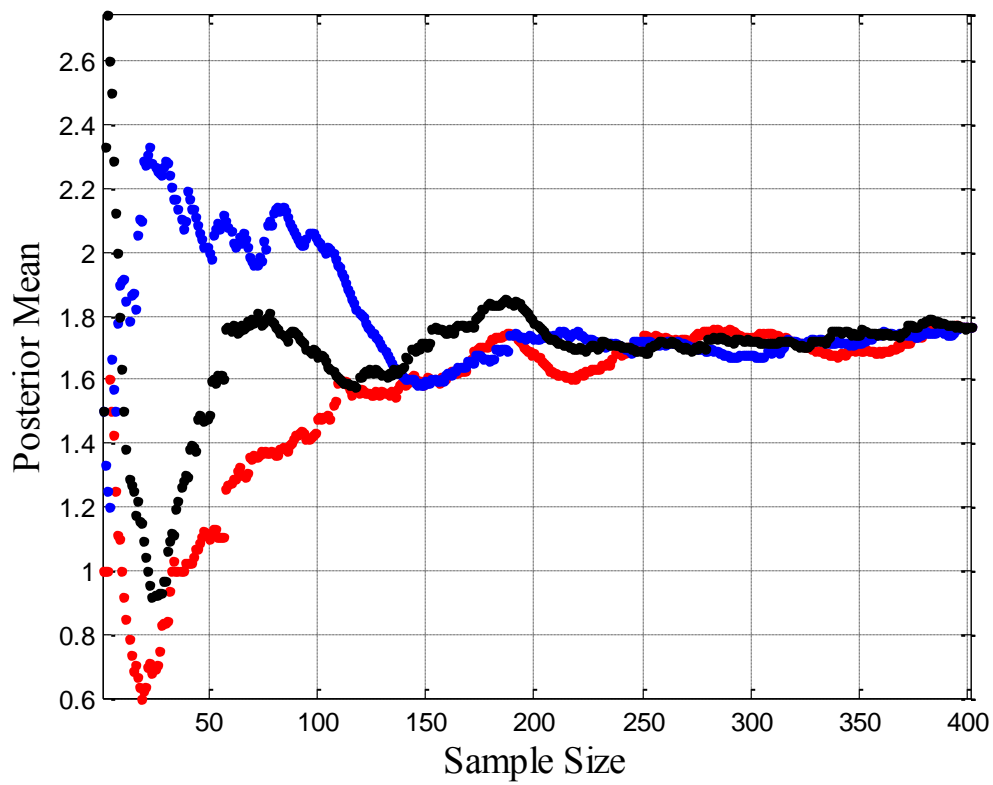
Figure 7. Covariate search results. The magenta dotted entries depict the log-marginal likelihood values derived from the search employing the Negative-Binomial generalized linear model with covariates #1, #2, ..., #13 included. The red dotted entries depict the log-marginal likelihood values derived from the search employing the Binomial generalized linear model with #1, #2, ..., #13 included. The magenta crossed entries depict the log-marginal likelihood values derived from the search employing the Negative-Binomial generalized linear model with covariates #1, #2, #3 included. The red crossed entries depict the log-marginal likelihood values derived from the search employing the Binomial generalized linear model with covariates #1, #2 and #3 included. The profiles from figure 5 have been superimposed in order to emphasize the differences obtained from moving between the conjugate- and generalized-linear-model specifications. The Negative Binomial model was iterated for 100 passes and the Binomial model was iterated for 200 passes. The predictive improvement, attributable to covariates, measured in terms of the log marginal likelihood differences, is substantial.

## **Supplemental Information**

### **Conjugate Analysis, Exchangeability and Prediction**

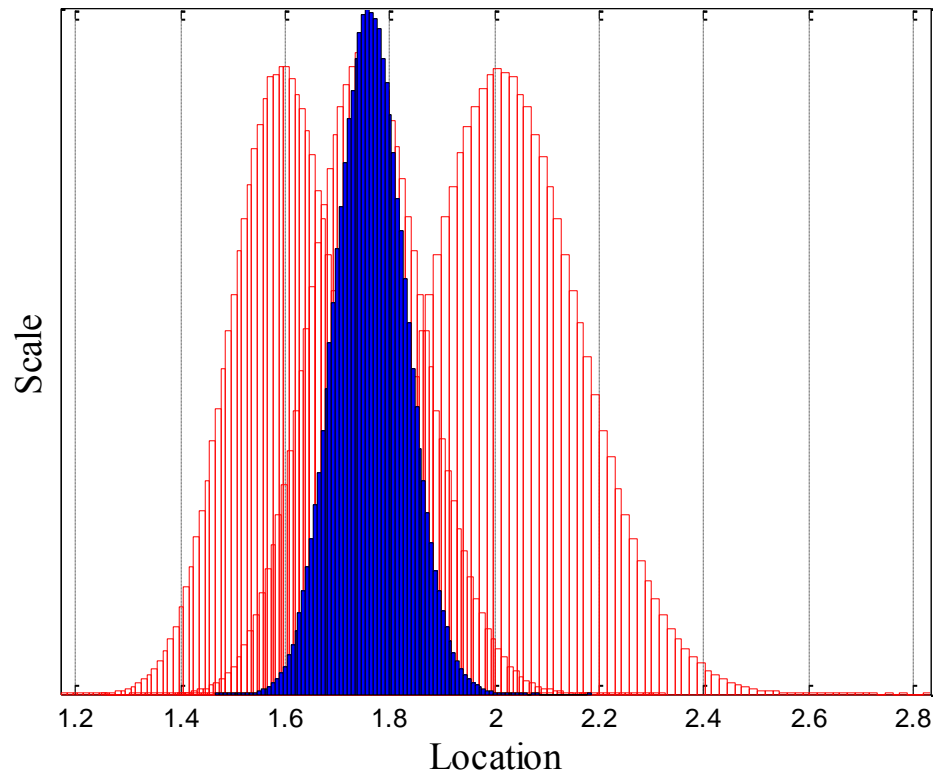
#### **S.1 Poisson Data Generation**

A first matter of some importance pertains to the way in which inferences are affected by the alternative production practices availed by the data across the dry-zone-, wet-zone- and intermediate-zone-climatic regions. Some alternatives for assessing this difference are available from exploiting the conjugate updating formulae which are availed by virtue of the conjugacy in the Poisson-Gamma setting. Supplemental Figure S.1.1 depicts the updating of the inferences about the Poisson parameter,  $\lambda$ , by sampling sequentially across the 401 observations, with the ordering of the sequence re-arranged so that, in turn, the dry-zone, then wet-zone and finally the intermediate-zone observations are placed first within the sample. There are some significant differences apparent across the climatic zones. These differences are further highlighted within the posterior reports of the distribution for  $\lambda$  across the respective climatic zones and the overall sample, as depicted in Supplemental Figure S.1.2. The predictive performance of the conjugate Poisson model is reported in Supplemental Figure S.1.3. A question concerning possible improvements in fit are considered by an extensive model selection exercise that considers, as datum, that differences across the climatic regimes exist but that adoption propensities may be affected by one or other of the covariates described in Table 1, justifications for which are noted in the text. These estimates are summarized in the degrees of fit depicted in Supplemental Figure S.1.4.

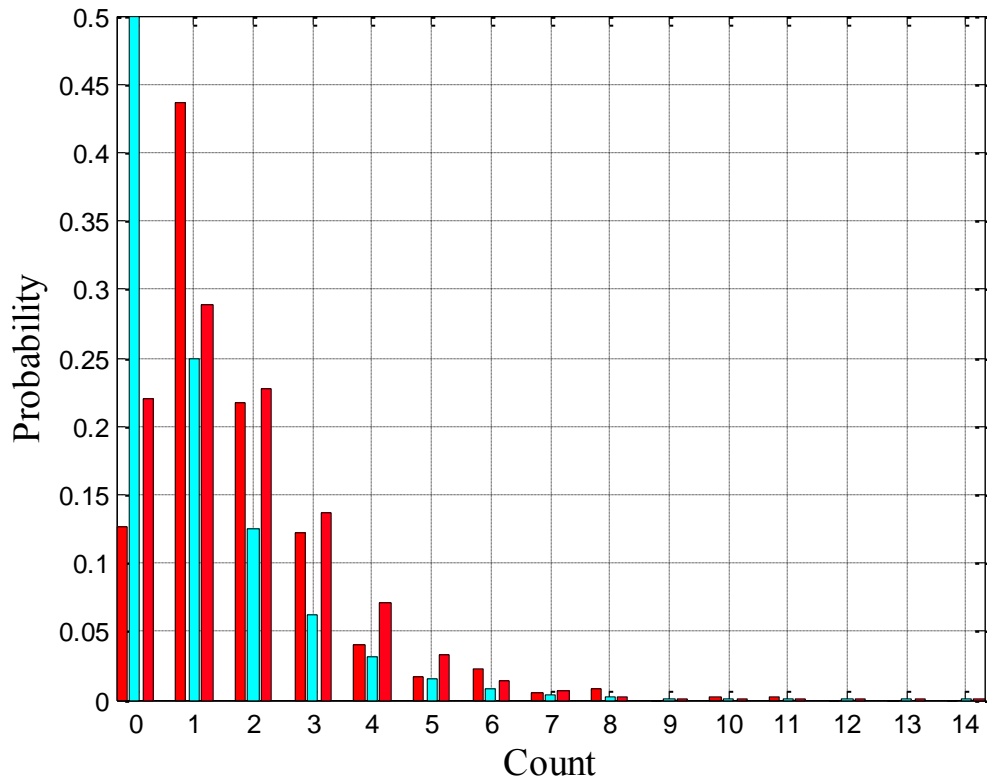


Supplemental Figure S.1.1. Sequential updating posterior expectations of the Poisson parameter. There are three, respective orderings. The blue dots depict the sequence when the dry-zone observations are sampled first; the red dots indicate the pattern of inferences when the wet-zone observations are sampled first; and the black dots indicate inference progression when the intermediate-zone observations are placed first.

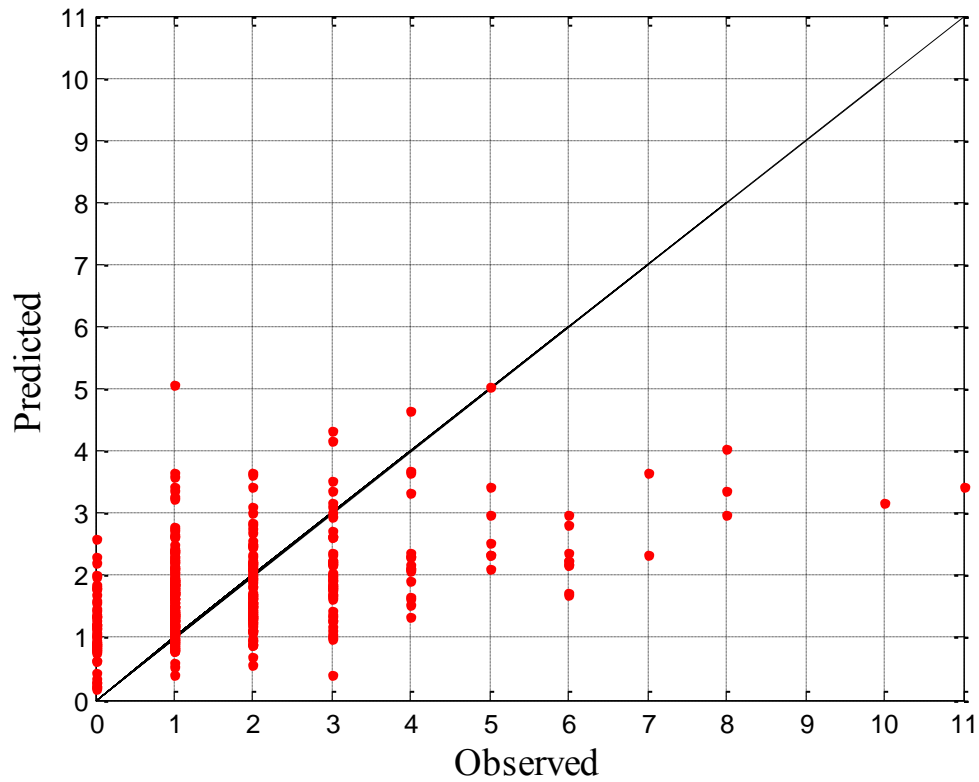




Supplemental Figure S.1.2. Posterior distributions of the Poisson parameter under alternative sampling assumptions. The blue-shaded area depicts the location and scale derived from processing the entire sample. The left-of-center distribution in red depicts the dry-zone inference; the right-of-center red distribution depicts the wet-zone inference; and the centered red distribution depicts the intermediate-zone inference.



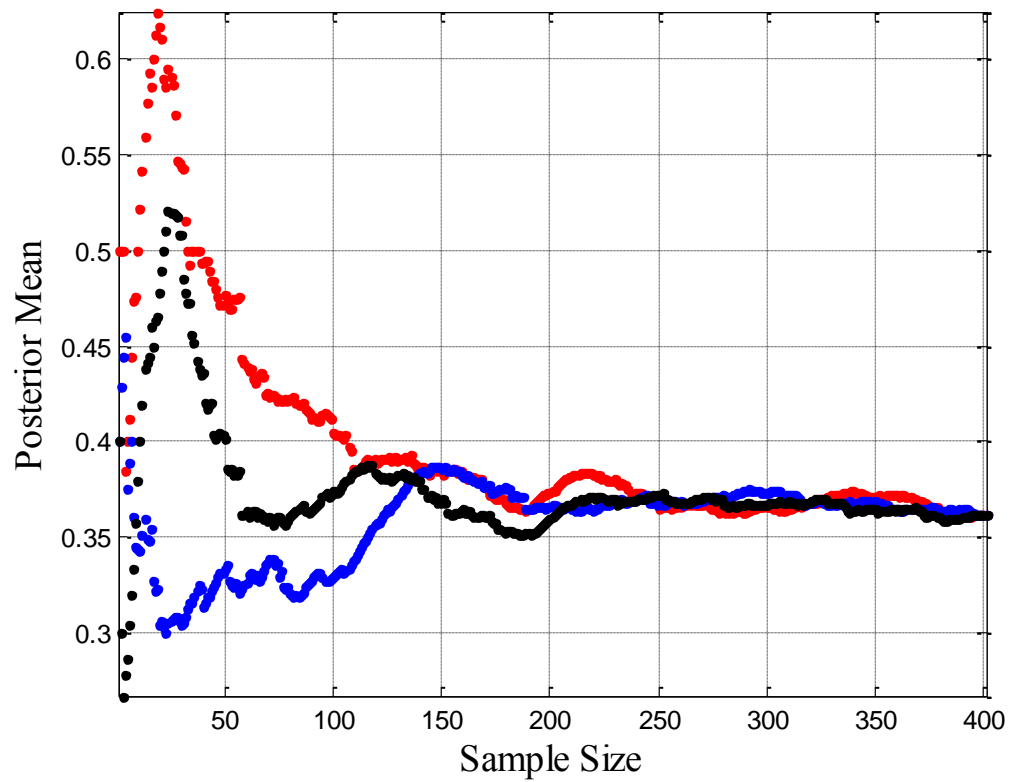
Supplemental Figure S.1.3. Posterior predictive inference for the counts derived from the Poisson sampling model. The right-most red-colored bars depict the observed data; the centered magenta-colored bars depict the predictions derived from the conjugate Poisson-Gamma model; and the left-most red-colored bars depict inferences derived from the preferred-covariate-specification of the generalized-linear Poisson regression model with covariates #1, #2, ..., #13 included.



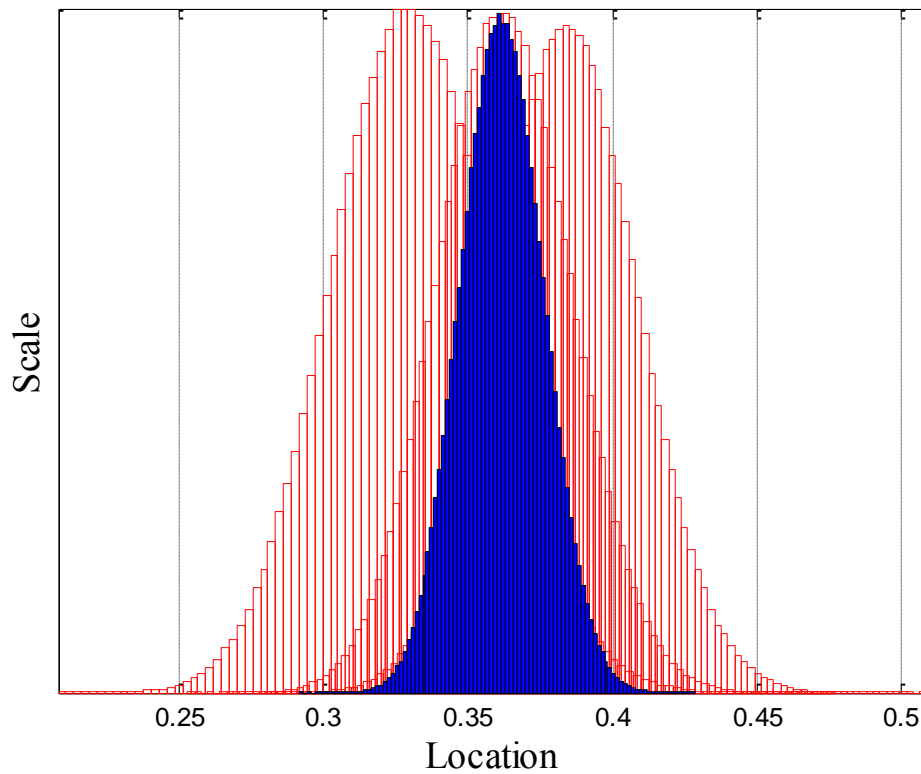
Supplemental Figure S.1.4. Posterior predictions derived from the preferred-covariate-specification of the generalized-linear Poisson regression model with covariates #1, #2, ..., #13 included. The line of perfect fit is denoted by black dots and the predicted observations derived from the Multivariate-Normal generalization of the Poisson-Gamma model are depicted in red. The coefficient of correlation between the observed and the predicted values is 0.49.

## **S.2 Geometric Data Generation**

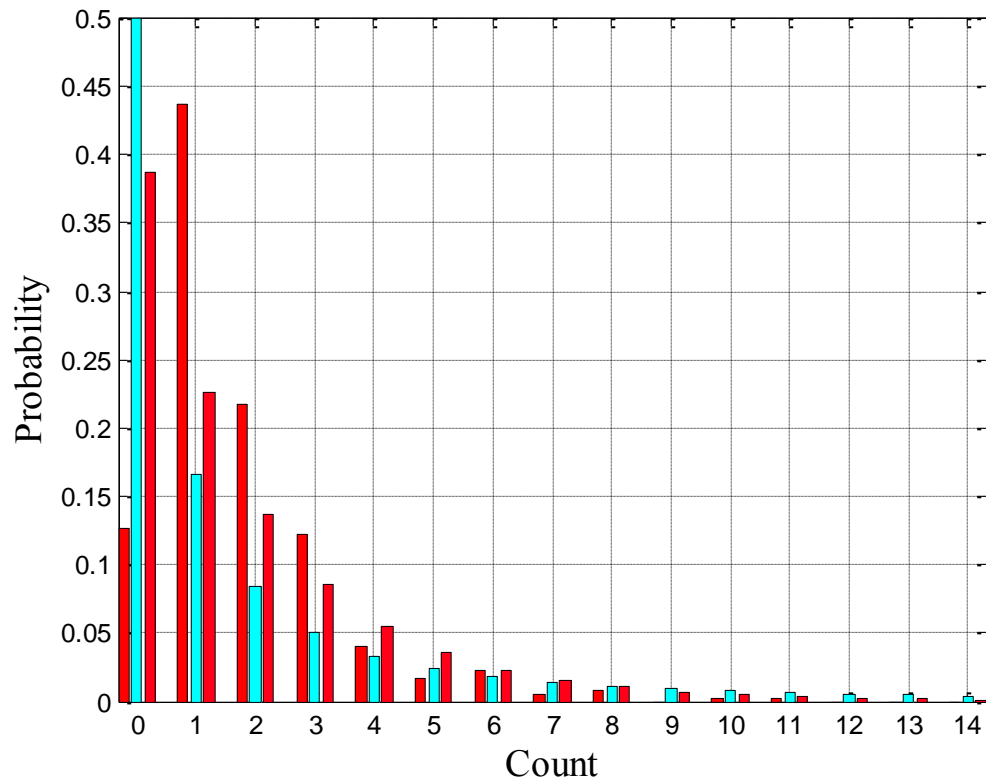
Once again, as we did for the Poisson model, we consider sequential updating and differences between the adoption potentials in the three distinct agro-climatic zones. Supplemental Figure S.2.1 depicts the inferences derived from sequentially updating the data ordered, respectively, with the dry-zone observations situated first in the sample; with the wet zone then situated first; and finally the intermediate zone observations situated first. The first ordering is depicted in red; the second in blue; and the third in black. Once again, some significant differences emerge. These differences are again manifested in posterior reports of the Geometric sampling model parameter. Supplemental Figure S.2.2 depicts these reports. Posterior predictive reports from the conjugate Geometric model are reported in Supplemental Figure S.2.3 and predictive reports corresponding to the generalized-linear regression estimation are reported in Supplemental Figure S.2.4.



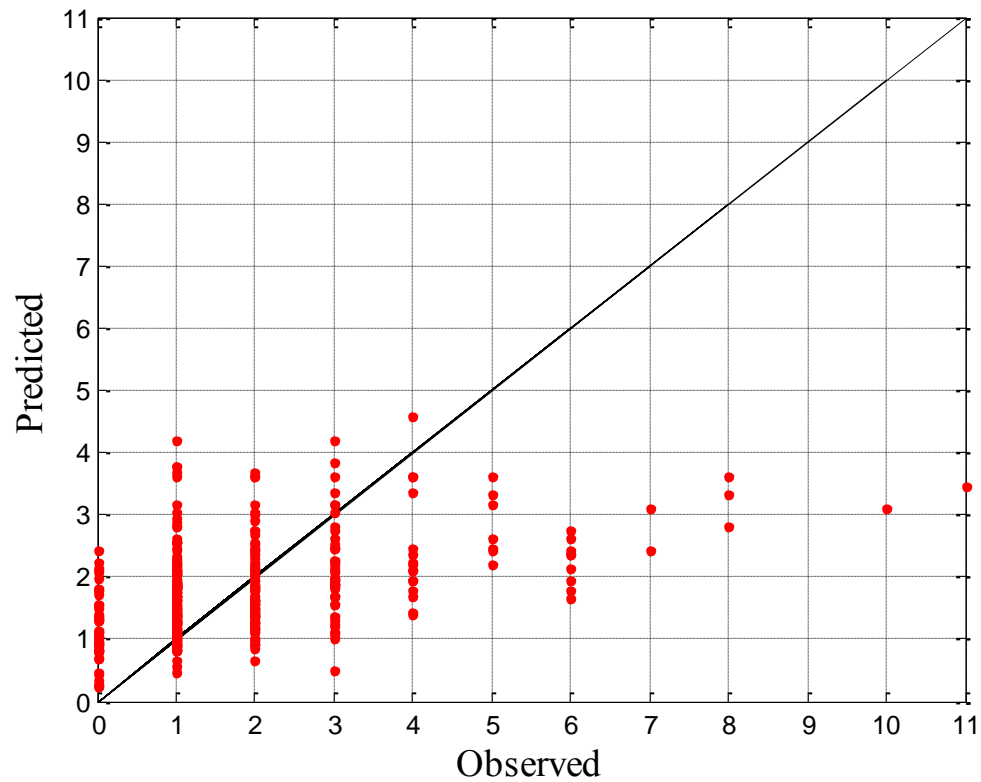
Supplemental Figure S.2.1. Sequential updating expectations about the Geometric parameter across the sample under three, respective data orderings. The blue dots depict the sequence when the dry-zone observations are sampled first; the red dots indicate the pattern of inferences when the wet-zone observations are sampled first; and the black dots indicate inference progression when the intermediate-zone observations are placed first.



Supplemental Figure S.2.2. Posterior distributions of the Geometric sampling parameter under alternative sampling assumptions. The blue-shaded area depicts the location and scale derived from processing the entire sample. The left-of-center distribution in red depicts the wet-zone inference; the right-of-center red distribution depicts the dry-zone inference; and the centered red distribution depicts the intermediate-zone inference.



Supplemental Figure S.2.3. Posterior predictive inference for the counts derived from the Geometric sampling model. The right-most red-colored bars depict the observed data; the centered magenta-colored bars depict the predictions derived from the conjugate Geometric-Beta model; and the left-most red-colored bars depict inferences derived from the preferred-covariate-specification of the generalized-linear Geometric regression model with covariates #1, #2, ..., #13 included.

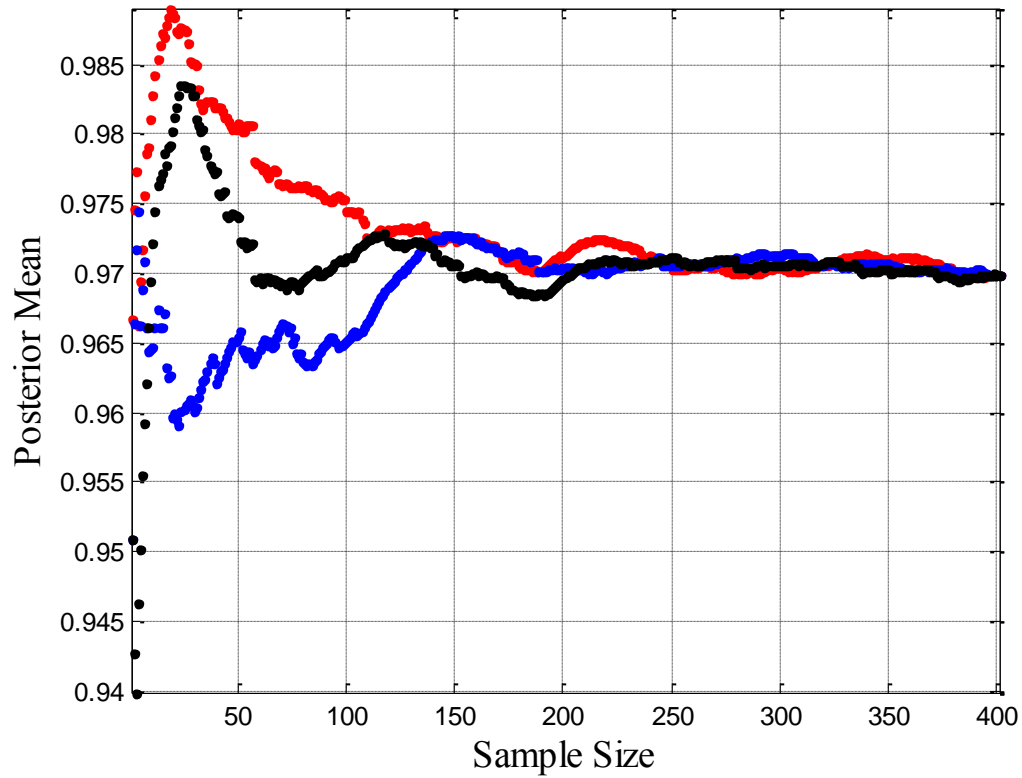


Supplemental Figure S.2.4. Posterior predictions derived from the preferred-covariate-specification of the generalized-linear Geometric regression model with covariates #1, #2, ..., #13 included. The line of perfect fit is denoted by black dots and the predicted observations derived from the Multivariate-Normal generalization of the Geometric-Beta model are depicted in red. The coefficient of correlation between the observed and the predicted values is 0.46.

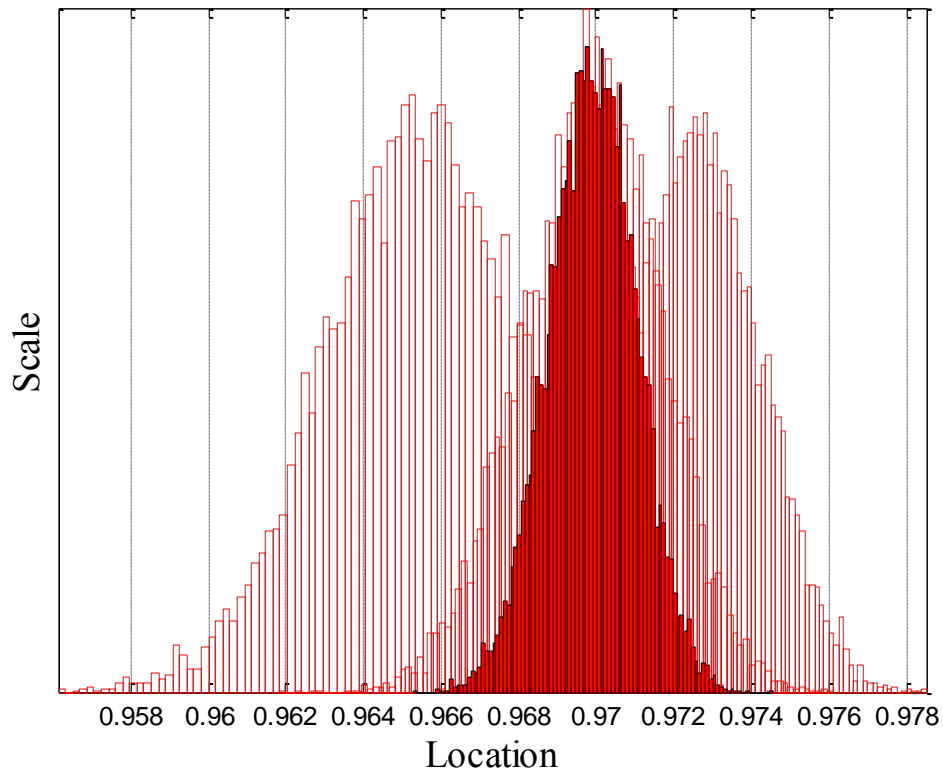


### **S.3 Negative Binomial Data Generation**

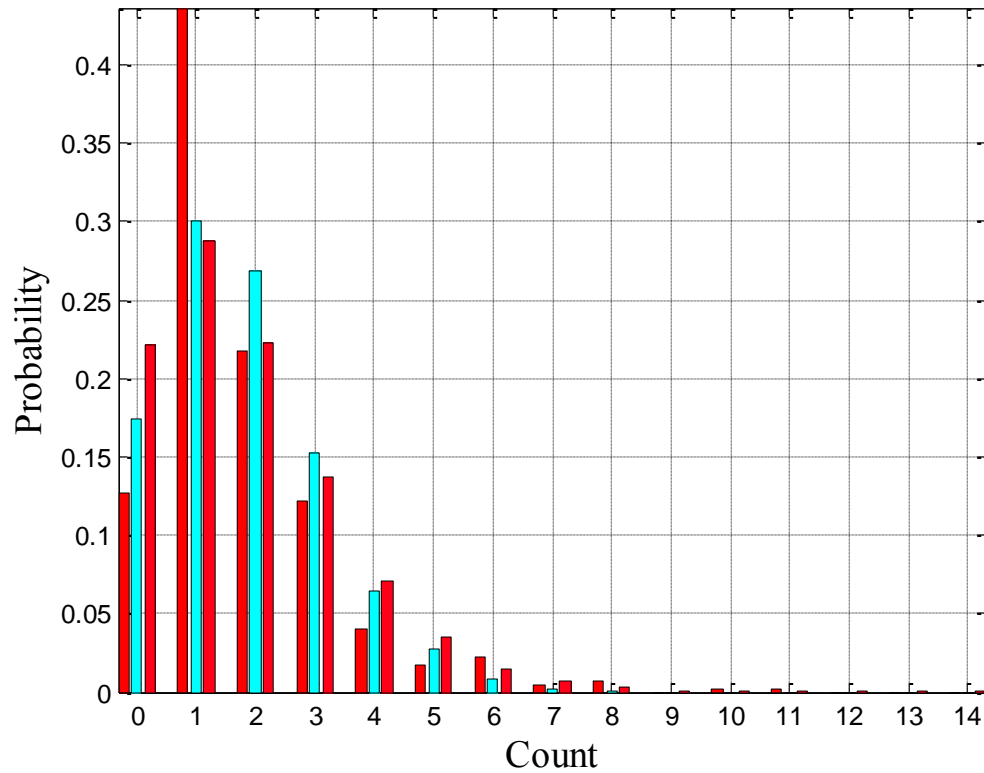
Once again, for the Negative Binomial model, we consider sequential updating and differences between the adoption potentials in the three distinct agro-climatic zones. Supplemental Figure S.3.1 depicts the inferences derived from sequentially updating the data ordered, respectively, with the dry-zone observations situated first in the sample; with the wet zone then situated first; and finally the intermediate zone observations situated first. The first ordering is depicted in red; the second in blue; and the third in black. Once again, some significant differences emerge. These differences are again manifested in posterior reports of the Negative Binomial sampling model parameter. Supplemental Figure S.3.2 depicts these reports. Posterior predictive reports from the conjugate Negative Binomial model are reported in Supplemental Figure S.3.3 and predictive reports corresponding to the generalized-linear regression estimation are reported in Supplemental Figure S.3.4.



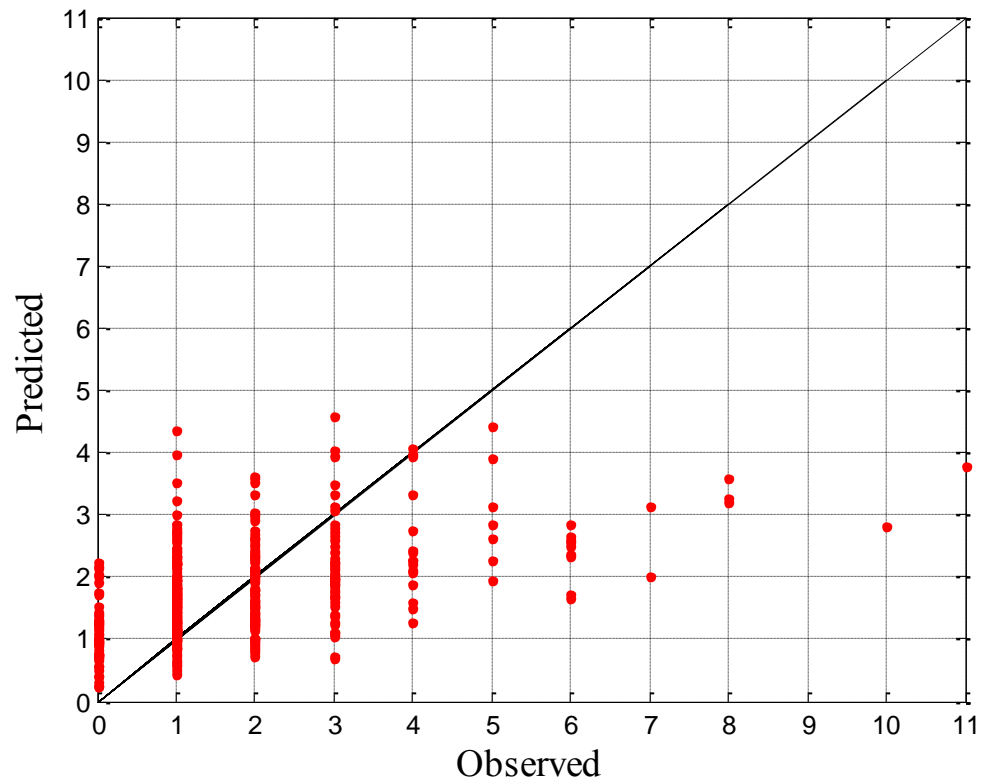
Supplemental Figure S.3.1. Sequential updating expectations about the Negative-Binomial parameter across the sample under three, respective data orderings. The blue dots depict the sequence when the dry-zone observations are sampled first; the red dots indicate the pattern of inferences when the wet-zone observations are sampled first; and the black dots indicate inference progression when the intermediate-zone observations are placed first.



Supplemental Figure S.3.2. Posterior distributions of the Geometric sampling parameter under alternative sampling assumptions. The blue-shaded area depicts the location and scale derived from processing the entire sample. The left-of-center distribution in red depicts the wet-zone inference; the right-of-center red distribution depicts the dry-zone inference; and the centered red distribution depicts the intermediate-zone inference.



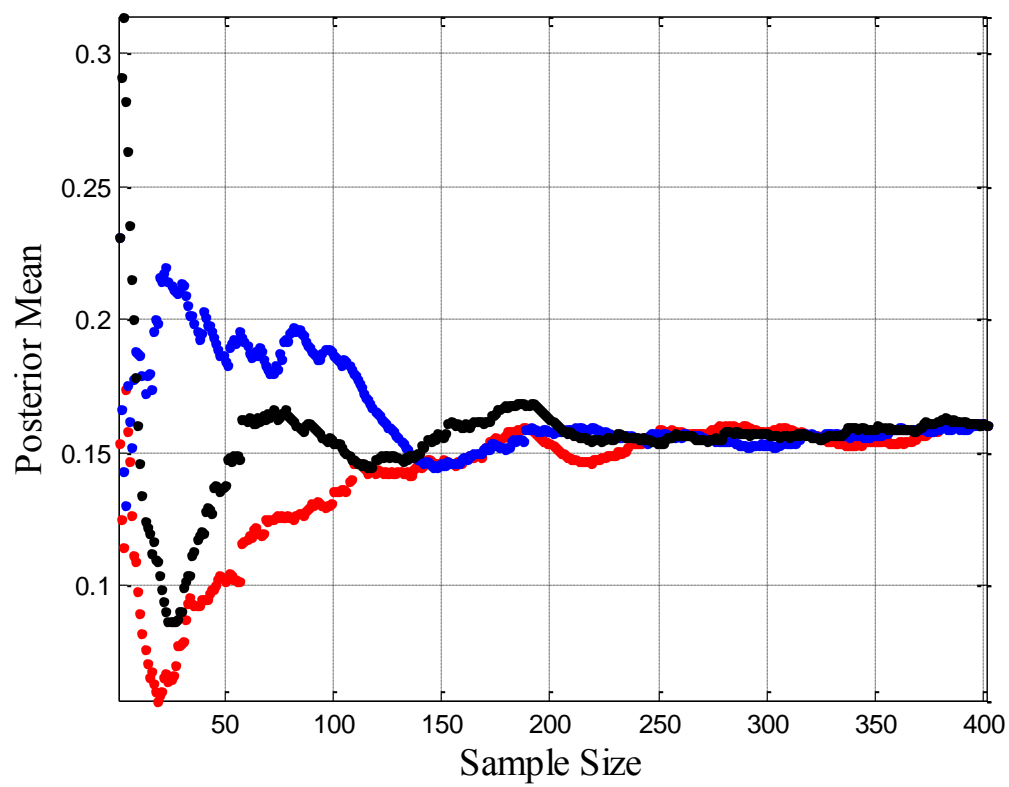
Supplemental Figure S.3.3. Posterior predictive inference for the counts derived from the Negative-Binomial sampling model. The right-most red-colored bars depict the observed data; the centered magenta-colored bars depict the predictions derived from the conjugate Negative-Binomial-Beta model; and the left-most red-colored bars depict inferences derived from the preferred-covariate-specification of the generalized-linear Negative Binomial regression model with covariates #1, #2, ..., #13, included.



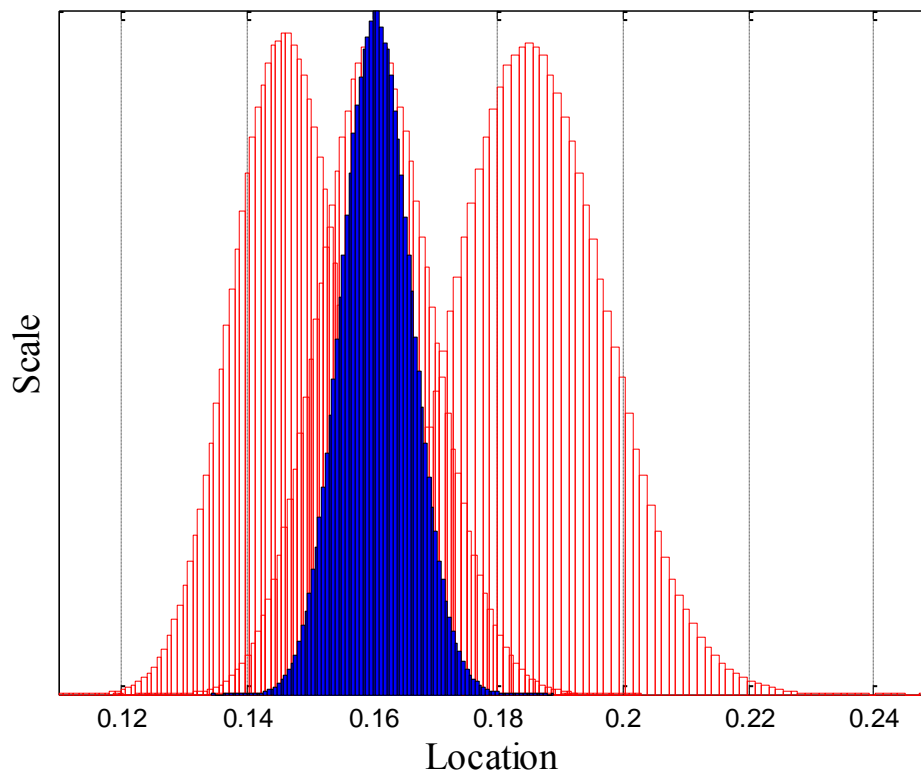
Supplemental Figure S.3.4. Posterior predictions derived from the preferred-covariate-specification of the generalized-linear Negative-Binomial regression model with covariates #1, #2, ..., #13, included. The line of perfect fit is denoted by black dots and the predicted observations derived from the Multivariate-Normal generalization of the Negative-Binomial-Beta model are depicted in red. The coefficient of correlation between the observed and the predicted values is 0.49.

## **S.4 Binomial Sampling Model**

Sequential updating and differences between the adoption potentials in the three distinct agro-climatic zones is undertaken for the Binomial sampling models. Supplemental Figure S.4.1 depicts the inferences derived from sequentially updating the data ordered, respectively, with the dry-zone observations situated first in the sample; with the wet zone then situated first; and finally the intermediate zone observations situated first. The first ordering is depicted in red; the second in blue; and the third in black. Significant differences emerge. These differences are again manifested in posterior reports of the Binomial sampling model parameter. Supplemental Figure S.4.2 depicts these reports. Posterior predictive reports from the conjugate Binomial model are reported in Supplemental Figure S.4.3 and predictive reports corresponding to the generalized-linear regression estimation are reported in Supplemental Figure S.4.4.

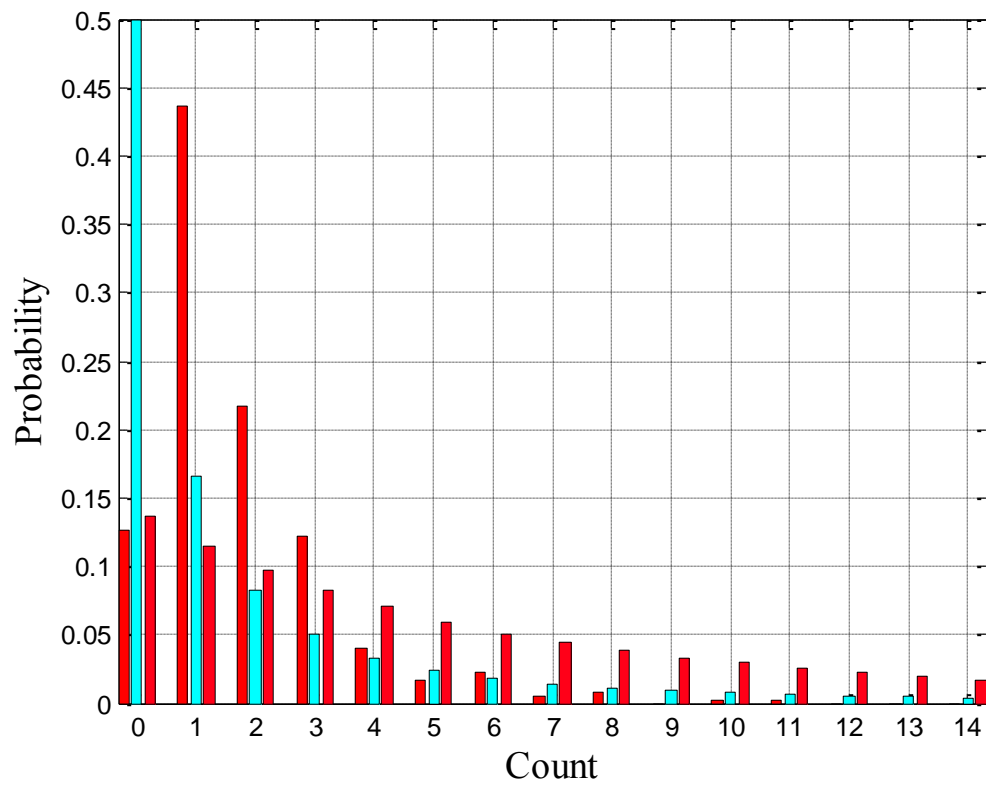


Supplemental Figure S.4.1. Sequential updating expectations about the Binomial parameter across the sample under three, respective data orderings. The blue dots depict the sequence when the dry-zone observations are sampled first; the red dots indicate the pattern of inferences when the wet-zone observations are sampled first; and the black dots indicate inference progression when the intermediate-zone observations are placed first.

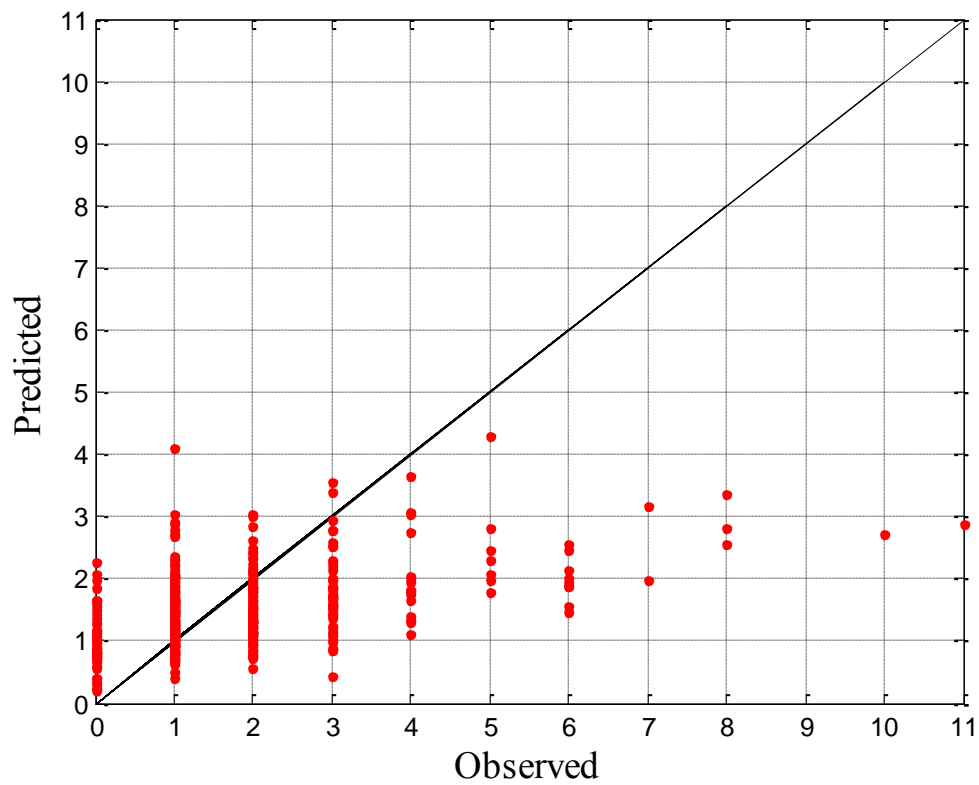


Supplemental Figure S.4.2. Posterior distributions of the Binomial sampling parameter under alternative sampling assumptions. The blue-shaded area depicts the location and scale derived from processing the entire sample. The left-of-center distribution in red depicts the dry-zone inference; the right-of-center red distribution depicts the wet-zone inference; and the centered red distribution depicts the intermediate-zone inference.





Supplemental Figure S.4.3. Posterior predictive inference for the counts derived from the Binomial sampling model. The right-most red-colored bars depict the observed data; the centered magenta-colored bars depict the predictions derived from the conjugate Geometric-Beta model; and the left-most red-colored bars depict inferences derived from the preferred-covariate-specification of the generalized-linear Poisson regression model with covariates #1, #2, ..., #13 included.



Supplemental Figure S.4.4. Posterior predictions derived from the preferred-covariate-specification of the generalized-linear Binomial regression model with covariates #1, #2, ..., #13, included. The line of perfect fit is denoted by black dots and the predicted observations derived from the Multivariate-Normal generalization of the Binomial-Beta model are depicted in red. The coefficient of correlation between the observed and the predicted values is 0.49.